

ROAD ENGINEERING FOR DEVELOPMENT

SECOND EDITION



Richard Robinson and Bent Thagesen

SPON
NOTES



**Also available as a printed book
see title verso for ISBN details**

Road Engineering for Development

Second edition

Developing countries in the tropics often have challenging natural conditions and different institutional and financial institutions than industrialized countries. However, most textbooks on road engineering are based on experience in industrialized countries with temperate climates, or deal only with specific issues.

Road Engineering for Development (published as *Highway and Traffic Engineering in Developing Countries* in its first edition) provides comprehensive coverage of the planning, design, construction and maintenance of roads in developing and emerging countries. It covers a wide range of technical and non-technical problems that may confront road engineers working in the developing world. This new edition has extended the focus to include those countries of the former Eastern Bloc, which share many institutional issues and the financial problems confronting countries from the South.

Designed as a fundamental text for civil engineering students, this book also offers a broad, practical view of the subject for practising engineers. It has been written with the assistance of a number of world-renowned specialist professional engineers with many years experience working in Africa, the Middle East, Central and Eastern Europe, Asia and Central America.

Dr Richard Robinson is an independent consultant and has an honorary appointment at the University of Birmingham in the United Kingdom. He spent over 20 years at the Transport Research Laboratory, and recently held the Senior Roads Specialist position at the European Bank for Reconstruction and Development. He has worked in 40 countries and has published over 100 papers and articles.

Professor Bent Thagesen is the former Professor of Highway Engineering at the Technical University of Denmark. He has worked as a highway engineer in Africa and in Asia, and has carried out research at the Danish Road Research Laboratory. He has been a consultant to the World Bank, the United Nations and various development organizations and firms concerned with roads in the developing world. He edited the first edition of this book *Highway and Traffic Engineering in Developing Countries*.

Road Engineering for Development

Second edition

Richard Robinson and
Bent Thagesen



Spon Press
Taylor & Francis Group

LONDON AND NEW YORK

First published 2004
by Spon Press
11 New Fetter Lane, London EC4P 4EE

This edition published in the Taylor & Francis e-Library, 2004.

“To purchase your own copy of this or any of Taylor & Francis or Routledge’s collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk.”

Spon Press is an imprint of the Taylor & Francis Group

© 2004 Richard Robinson and Bent Thagesen

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

Road engineering for development/edited by Richard Robinson and Bent Thagesen. – 2nd ed.

p. cm.

Rev. ed. of: Highway and traffic engineering in developing countries. 1995.

Includes bibliographical references and index.

I. Highway engineering—Developing countries. I. Robinson, Richard, 1946—
II. Thagesen, Bent. III. Title: Highway and traffic engineering in developing countries.

TE127.R63 2003

625.7'09172'4—dc21

2003000525

ISBN 0-203-30198-6 Master e-book ISBN

ISBN 0-203-34045-0 (Adobe eReader Format)

ISBN 0-415-27948-8 (hbk)

ISBN 0-415-31882-3 (pbk)

Contents

<i>List of illustrations</i>	xii
<i>List of contributors</i>	xix
<i>Preface</i>	xxi
<i>Acknowledgements</i>	xxiii

PART I

Planning

I

1 Roads and development

3

PETER BROCH, BRITHA MIKKELSEN AND RICHARD ROBINSON

- 1.1 The meaning of development 3*
- 1.2 Evolution of development theory 6*
- 1.3 Poverty 10*
- 1.4 Evolution of road development 13*
- 1.5 Impact of roads on economic development 14*
- 1.6 Transport and poverty 15*
- 1.7 Impact of roads on the transition to a market economy 20*
- References 21*

2 Policy

23

RICHARD ROBINSON

- 2.1 The nature of policy 23*
- 2.2 Government policy 24*
- 2.3 Organizational policy 26*
- 2.4 Integrated policy 28*
- 2.5 Dissemination 31*
- 2.6 Performance indicators 31*
- 2.7 Policy formulation 33*
- 2.8 Policy issues 36*
- References 37*

3	Traffic	38
	KENT FALCK-JENSEN, JAN KILDEBOGAARD AND RICHARD ROBINSON	
	3.1 <i>Introduction</i>	38
	3.2 <i>Traffic composition</i>	38
	3.3 <i>Traffic flows and growth</i>	42
	3.4 <i>Capacity and speed-flow</i>	43
	3.5 <i>Increasing capacity</i>	50
	3.6 <i>Traffic information and data</i>	51
	<i>References</i>	56
4	Traffic safety	57
	GOFF JACOBS AND CHRIS BAGULEY	
	4.1 <i>Background</i>	57
	4.2 <i>Nature of the problem</i>	57
	4.3 <i>Road accident costs</i>	59
	4.4 <i>Contributory factors</i>	62
	4.5 <i>Institutions and information systems</i>	64
	4.6 <i>Improving road safety</i>	65
	4.7 <i>Priorities for action</i>	75
	<i>References</i>	76
5	Roads and the environment	78
	WILLIAM V. KENNEDY	
	5.1 <i>Introduction</i>	78
	5.2 <i>Environmental impacts and their mitigation</i>	78
	5.3 <i>Environmental impact assessment</i>	86
	5.4 <i>Concluding remarks</i>	90
	<i>References</i>	91
6	Planning methods	92
	KNUD RASK OVERGAARD	
	6.1 <i>The nature of planning</i>	92
	6.2 <i>Strategic planning</i>	93
	6.3 <i>Physical network planning</i>	95
	6.4 <i>Transport demand forecasting</i>	98
	6.5 <i>Plan development and implementation</i>	106
	6.6 <i>Planning for rural transport infrastructure</i>	110
	<i>References</i>	112

7 Economic appraisal	114
RICHARD ROBINSON	
7.1 <i>Purpose</i>	114
7.2 <i>Preliminary consideration</i>	114
7.3 <i>Cost estimation</i>	117
7.4 <i>Assessment of benefits for major roads</i>	119
7.5 <i>Cost–benefit analysis for major roads</i>	122
7.6 <i>Minor road appraisal</i>	127
<i>References</i>	132
PART II	
Design	135
8 Soil investigation	137
BENT THAGESEN	
8.1 <i>Introduction</i>	137
8.2 <i>Evaluation of existing information</i>	138
8.3 <i>Field investigation</i>	144
8.4 <i>Laboratory testing</i>	147
8.5 <i>Soil classification</i>	157
<i>References</i>	161
9 Tropical soils and rocks	162
BENT THAGESEN	
9.1 <i>Introduction</i>	162
9.2 <i>Rocks</i>	163
9.3 <i>Soils</i>	165
<i>References</i>	177
10 Hydrology and drainage	178
TUE HALD, JAN HASSING, MICHAEL HØGEDAL AND ALAN JACOBSEN	
10.1 <i>Introduction</i>	178
10.2 <i>Rainfall</i>	179
10.3 <i>Flood discharge estimation</i>	183
10.4 <i>Hydraulic design</i>	188
10.5 <i>Longitudinal drainage components</i>	190
10.6 <i>Cross drainage components</i>	193
10.7 <i>Culvert design</i>	196
10.8 <i>Erosion and scour protection</i>	199
<i>References</i>	204

11 Geometric design controls	205
KENT FALCK-JENSEN	
11.1 <i>Basic considerations</i>	205
11.2 <i>Approach to selecting design standards</i>	207
11.3 <i>Classification of roads</i>	208
11.4 <i>Sight distance</i>	211
11.5 <i>Traffic</i>	215
11.6 <i>Cross-section</i>	215
<i>References</i>	223
12 Geometric alignment design	224
KENT FALCK-JENSEN	
12.1 <i>Geometric design elements</i>	224
12.2 <i>Horizontal alignment</i>	225
12.3 <i>Vertical alignment</i>	229
12.4 <i>Phasing</i>	234
12.5 <i>Alignment selection</i>	235
12.6 <i>Intersections</i>	239
12.7 <i>Low-cost roads</i>	241
12.8 <i>Computer-aided design</i>	242
<i>References</i>	243
13 Earthworks, unbound and stabilized pavements	244
BENT THAGESEN	
13.1 <i>Earthworks</i>	244
13.2 <i>Pavement structure</i>	251
13.3 <i>Unbound pavement layers</i>	253
13.4 <i>Design of gravel pavements</i>	256
13.5 <i>Stabilized pavement layers</i>	257
<i>References</i>	263
14 Asphalt pavement materials	264
RICHARD KOOLE AND BENT THAGESEN	
14.1 <i>Asphalt pavements</i>	264
14.2 <i>Bituminous binders</i>	264
14.3 <i>Surface dressing</i>	268
14.4 <i>Premixed asphalt</i>	273
<i>References</i>	283

15 Structural design of asphalt pavements	284
JOHN ROLT	
15.1 <i>Introduction</i>	284
15.2 <i>Basic empirical methods</i>	285
15.3 <i>Overseas Road Note 31</i>	286
15.4 <i>The AASHTO method</i>	289
15.5 <i>Theoretical-mechanistic design</i>	295
15.6 <i>Overlay design</i>	302
<i>References</i>	304
PART III	
Construction	307
16 Contracts and works procurement	309
TORBEN MIKKELSEN AND TIM WAAGE	
16.1 <i>Project execution methods</i>	309
16.2 <i>Types of contract</i>	313
16.3 <i>The FIDIC contract</i>	316
<i>References</i>	323
17 Contract supervision	325
TORBEN MIKKELSEN, BENT THAGESEN AND TIM WAAGE	
17.1 <i>Introduction</i>	325
17.2 <i>Supervision organization</i>	325
17.3 <i>Quality control</i>	329
17.4 <i>Measurement of work</i>	332
17.5 <i>Payment to the contractor</i>	333
17.6 <i>Progress control</i>	336
17.7 <i>Extension of time</i>	337
17.8 <i>Cost claims</i>	338
17.9 <i>Default of contractor</i>	341
17.10 <i>Supervision procedures</i>	342
<i>References</i>	344
18 Appropriate technology	345
JAN DE VEEN, RICHARD ROBINSON AND BENT THAGESEN	
18.1 <i>Technology options</i>	345
18.2 <i>Choice of technology</i>	345
18.3 <i>Intermediate methods</i>	352
18.4 <i>Equipment management</i>	364
<i>References</i>	367

PART IV	
Maintenance	369
19 Maintenance management	371
RICHARD ROBINSON	
19.1 <i>The road network as an asset</i>	371
19.2 <i>Road network management</i>	371
19.3 <i>Network information</i>	374
19.4 <i>Assessing needs</i>	376
19.5 <i>Determining options</i>	381
19.6 <i>Choosing actions</i>	383
19.7 <i>Implementation</i>	386
19.8 <i>Monitoring and audit</i>	389
19.9 <i>Information systems</i>	390
References	392
20 Maintenance operations	393
BENT THAGESEN	
20.1 <i>Introduction</i>	393
20.2 <i>Classification of maintenance activities</i>	393
20.3 <i>Safety</i>	394
20.4 <i>Asphalt pavements</i>	395
20.5 <i>Unpaved roads</i>	402
20.6 <i>Roadside areas</i>	407
20.7 <i>Drainage systems</i>	409
20.8 <i>Traffic control devices</i>	414
20.9 <i>Implementation</i>	414
References	415
21 The HDM-4 road investment model	416
HENRY KERALI	
21.1 <i>Introduction</i>	416
21.2 <i>Types of analysis</i>	417
21.3 <i>Structure of HDM-4</i>	420
21.4 <i>Components of HDM-4</i>	423
21.5 <i>Examples of applications</i>	431
References	437

PART V	
Institutional issues	439
22 Institutional development	441
RICHARD ROBINSON	
22.1 <i>Introduction</i>	441
22.2 <i>Finance</i>	443
22.3 <i>Management</i>	450
22.4 <i>Ownership and responsibility</i>	457
22.5 <i>Expected outcomes</i>	458
<i>References</i>	458
23 Training of staff	460
RICHARD ROBINSON AND BENT THAGESEN	
23.1 <i>Purpose of training</i>	460
23.2 <i>Institutional issues</i>	462
23.3 <i>Training types</i>	463
23.4 <i>Training needs analysis</i>	466
23.5 <i>Planning</i>	468
23.6 <i>Detailed preparation</i>	475
23.7 <i>Implementation</i>	479
<i>References</i>	482
24 Development aid	483
JENS ERIK BENDIX RASMUSSEN AND TONNY BAEK	
24.1 <i>Introduction</i>	483
24.2 <i>Resource transfers</i>	483
24.3 <i>International aid agencies</i>	487
24.4 <i>Bilateral donors</i>	489
24.5 <i>Assistance to the transport and road sector</i>	490
24.6 <i>Recent trends in aid management</i>	491
24.7 <i>Project cycle and framework</i>	494
<i>References</i>	500
<i>Index</i>	502

Illustrations

Figures

1.1	Relationship between GNI per capita and HDI	6
2.1	Transport policy model	25
2.2	Relationship between components of the policy framework	27
2.3	Example of relationship between policies produced at different levels of government	28
2.4	Example of national policy statement and organizational policy framework	30
2.5	Example of a pie chart presentation of a performance indicator showing proportions of road network in different conditions	33
2.6	Example of a histogram presentation of a performance indicator showing achievement of pot-hole patching over time	33
3.1	Non-motorized transport	40
3.2	Relationship between equivalence factor and axle loading	44
3.3	Speed–flow diagram	47
3.4	Relationship between daily variability and traffic flow	55
3.5	Errors in AADT estimates from random counts of varying duration	55
4.1	Fatalities/10,000 licensed motor vehicles in selected countries	58
4.2	Regional trends in road deaths 1980–95	59
4.3	Pedestrian fatalities as a percentage of all road accident fatalities	60
4.4	Examples of effects of engineering design on road safety	68
5.1	Comparison of water impact mitigation measures	80
5.2	Comparison of flora and fauna protection mitigation measures	81
5.3	Comparison of erosion mitigation measures	82
5.4	Comparison of noise mitigation measures	85
6.1	Example of budget required to maintain a given road condition	94
6.2	Example of projected road conditions resulting from the application of a fixed budget	94
6.3	Steps in strategic planning for two examples of decision	95
6.4	Example of main activity areas and major transport desire lines	97
6.5	Example of existing and potential new road links	98
6.6	Optimum road standard	99
6.7	Traffic development on an improved road	99
6.8	Modal split model	104

6.9	Methodology for passenger traffic forecasting and system evaluation	105
6.10	Effect of equity objective on economic efficiency standard	107
6.11	Roads investment programme	108
7.1	Use of a demand curve to assess generated traffic benefits	122
7.2	Annual cycle of cost and road deterioration	124
7.3	Indicative VOCs for different roughness values	132
8.1	Dendritic and trellis drainage pattern	140
8.2	Typical gully cross-sections	140
8.3	Tropical and subtropical areas covered by terrain classification	143
8.4	Block diagram of a land system with five land facets	144
8.5	Equipotentials and current lines in a homogeneous earth	145
8.6	Quartering	147
8.7	Particle-size distribution curves	148
8.8	Drive cylinder	150
8.9	Sand cone	151
8.10	Example of compaction curves	152
8.11	Equipment for CBR testing	153
8.12	Los angeles abrasion machine	157
8.13	Unified classification system	158
9.1	Distribution of lateritic soils	166
9.2	Typical gradings of lateritic gravels	168
9.3	Relation between water absorption and los angeles abrasion value for typical West African lateritic gravels	169
9.4	Distribution of expansive clays in Africa	172
9.5	Potential swelling tendencies of clays	174
10.1	Monthly rainfall distribution in Lusaka, Zambia, 1974–2000	180
10.2	Area reduction for a 5-year recurrence interval	181
10.3	Extreme value probability distribution function	183
10.4	Intensity–duration curves	184
10.5	A flood hydrograph combining flows from several mountain catchments	185
10.6	Chute outlet	192
10.7	Drift	194
10.8	Pipe culvert	195
10.9	Types of culvert outlets	196
10.10	Inlet control	198
10.11	Outlet control	199
10.12	Ditch checks	201
10.13	Shields diagram	204
11.1	Design process	209
11.2	Typical cross-sections	217
11.3	Superelevation design curves	221
12.1	Minimum offset to obstruction adjacent to curve	228
12.2	Speed reduction effect of upgrade	231
12.3	Length of crest vertical curves for safe stopping sight distance	233
12.4	Length of sag vertical curve for adequate riding comfort	235

12.5	Factors affecting road alignment selection	238
12.6	Dimensions for typical priority T-junction (driving on the left)	241
13.1	Bulldozer	245
13.2	Front-end loader	246
13.3	Motor grader	247
13.4	Cross-section of a road in steep side-sloping terrain	248
13.5	Self-propelled smooth steel-wheeled roller	250
13.6	Relationship between relative compaction of a sandy clay and the number of passes of smooth steel-wheeled rollers	251
13.7	Terms used for the different pavement layers	252
13.8	Effect of the addition of lime on London Clay	260
14.1	Penetration test	265
14.2	Softening point test	265
14.3	Choice of binder for surface dressing	270
14.4	Rate of spread of binder for 3,000-grade cutback	272
14.5	Marshall testing machine	278
14.6	Results from a marshall test	279
14.7	Batch mixer	280
14.8	Drum mixer	281
14.9	Asphalt paver	283
15.1	Key to the design charts in <i>ORN31</i>	290
15.2	Design chart from <i>ORN31</i>	291
15.3	Design chart from <i>ORN31</i>	292
15.4	Falling weight deflectometer	298
15.5	Effect of temperature on the elastic modulus of asphalt mixes	299
15.6	Fatigue laws for pavement materials	299
15.7	Shifting the fatigue line	300
15.8	The effect of ageing on fatigue	300
17.1	Example of a road supervision organization	328
18.1	Comparison of cost between equipment and labour	348
18.2	Comparison of productivity between equipment and labour	349
18.3	Decision-making regarding technology choice	351
18.4	Gravelling rural access roads in Kenya	356
18.5	Digging of ditches	357
18.6	Head baskets	358
18.7	Appropriate design of wheelbarrow	359
18.8	Improved animal harness	360
18.9	Available earthmoving methods	361
19.1	Example of the management cycle for programming	374
19.2	Inspection form for paved roads	379
19.3	Deformation gauge	380
19.4	BI unit for measuring roughness	381
19.5	Example of use of condition indices for treatment selection	383
19.6	Example of worksheet	389
20.1	Rutting	397
20.2	Edge subsidence	397
20.3	Shoving	398

20.4	Broken edges	399
20.5	Patching of pot-holes	400
20.6	Initial deterioration – rutting without shoving	402
20.7	Beam drag and tyre drag	404
20.8	Hand-loading of trailers	406
20.9	Outfall basin	413
21.1	HDM-4 system structure	420
21.2	Life cycle analysis using HDM-4	423
21.3	The conceptual impact over time of maintenance treatments on pavement condition (roughness) relative to a maintenance standard	425
21.4	Effect of road condition on vehicle operating costs for rolling terrain	429
21.5	HDM-4 output for programme analysis when budget is unconstrained	434
21.6	HDM-4 output for programme analysis with optimized work programme when budget is constrained	435
22.1	Transition of road financing over time	449
22.2	Model road network management arrangements	454
22.3	Five-stage strategy for restructuring the organization of road network management in Romania	455
23.1	Staffing and skills inventory form	467
23.2	Job description for a district maintenance engineer in Nigeria	468
23.3	School questionnaire	469
23.4	Training programme from Sabah, Malaysia	472
23.5	Description of a course for senior field officers in Zimbabwe	473
23.6	Lesson plan for a training course for road engineers in the Peoples Republic of China	476
23.7	Course note from a training course for road overseers in Sabah, Malaysia	478
24.1	Net ODA transfers	484
24.2	Net flow of ODA funds 1993–2000	485
24.3	The project cycle	496

Tables

1.1	Summary data on GNI and population growth by region	5
1.2	Road densities for selected countries	15
3.1	Example of a traffic classification	39
3.2	The distribution of motor vehicles in some selected countries	39
3.3	Basic vehicle types and their performance characteristics	40
3.4	Public transport distribution in selected cities	41
3.5	Passenger car equivalent (<i>E</i>) by road and carriageway type for Jakarta	45
3.6	Typical PCSE values for different road types	45
3.7	Variations in 30th HV	46
3.8	Level of service for rural road design	47

3.9	Volume–capacity ratio	48
3.10	Lane width factor, f_W	49
3.11	Passenger car equivalents	49
3.12	Directional factor, f_D	49
3.13	Maximum AADT for two-lane rural roads	50
3.14	Comparison of the capacity of an approach with mixed respectively separated traffic	51
3.15	Types of urban traffic surveys	52
3.16	Traffic volume and axle loading data for different IQL values	53
3.17	Examples of modal share for motorized trips only	54
4.1	Characteristics of fatal accidents	60
4.2	Road accident costs by region 1997	62
4.3	Causes of road accidents as determined by the police in different countries	63
6.1	Matrix for trips between n zones	101
8.1	AASHTO soil classification system	160
9.1	Classification of igneous rocks	163
10.1	Run-off coefficient for the rational method	188
10.2	Recommended design standards for road classes and drainage components	189
10.3	Probability of design standards being exceeded during design life	190
10.4	Manning roughness coefficient	193
10.5	Permissible velocities for ditches lined with various grass covers	200
10.6	Permissible velocities in excavated ditches	200
11.1	Road design standards	210
11.2	Design speed and design class	211
11.3	Coefficient of longitudinal friction, f	213
11.4	Stopping sight distance on level grade	213
11.5	Stopping sight distance for different gradients	213
11.6	Passing sight distances	214
11.7	Pavement, shoulder and formation widths of rural roads in Bhutan	218
11.8	Standard cross-sections for rural roads	218
11.9	Curve widening (TRRL)	219
11.10	Curve widening (SATCC)	219
11.11	Maximum and minimum values for the relative longitudinal gradient	222
11.12	Suggested earth slopes for design	222
12.1	Minimum radii for horizontal curves in metres	226
12.2	Minimum radii of curve with adverse crossfall in metres	227
12.3	Maximum recommended gradient	230
12.4	Maximum gradients depending on type of terrain	230
12.5	Minimum levels of acceptable vertical acceleration	234
12.6	Intersection layout standards	241
13.1	Recommended particle size distribution for granular sub-base materials	254
13.2	Maximum plasticity values for sub-base materials	254

13.3	Recommended particle size distribution for natural gravel base materials	255
13.4	Recommended particle size distribution for crushed gravel and crushed rock base materials	256
13.5	Recommended strength requirements for cement and lime stabilized sub-base and base materials	259
14.1	Recommended nominal chipping size in mm for surface dressing	269
14.2	Condition factors for determining the rate of application of binder for surface dressing	271
14.3	Spraying temperatures for binders	273
14.4	Requirements for aggregates for hot-mix asphalt	275
14.5	Recommended particle size distribution for combined aggregate for use in asphalt concrete	276
14.6	Recommended marshall design criteria for asphalt concrete wearing course	280
15.1	Estimated minimum design CBR values in <i>ORN31</i>	289
15.2	Typical strength coefficients	293
15.3	Typical values of elastic modulus for pavement materials	296
19.1	Example of some intervention levels for paved roads	382
19.2	Examples of work outputs from activities	384
20.1	Activities included in the maintenance of the drainage system	410
22.1	Summary of road network management models for administrative decentralization	456
24.1	ODA for selected countries	485
24.2	Net resource flows from DAC member states to the developing countries	486
24.3	ODA from the DAC countries	486
24.4	Net disbursements from financial institutions	488
24.5	Net disbursements from UN agencies	489
24.6	ODF to transport and communications	491

Boxes

1.1	Development	3
1.2	Economic and statistical concepts	4
1.3	Community and individual poverty	12
1.4	Selected recommendations in support of social sustainability	19
2.1	Extract from a government policy statement from New Zealand	24
2.2	Extract from a draft government policy statement from Zimbabwe	26
2.3	Performance indicators and operational statistics	31
2.4	Examples of performance monitoring against policy objectives	32
2.5	Extracts from performance indicators published by a road administration	33
2.6	Case study of policy formulation	34
2.7	Key issues for policy and its development	35
2.8	Examples of road sector policy issues	36
4.1	Accident costing methods	61

4.2	Microcomputer Accident Analysis Package	66
4.3	Selected engineering and planning measures for accident prevention	67
4.4	Road safety audit roles and responsibilities	69
4.5	Countermeasures	71
4.6	Type and dimensions of road hump	74
5.1	Typical components of an EIA	88
6.1	Integrated rural accessibility planning	112
7.1	Major and minor roads	115
7.2	Access and level of service	127
7.3	Example of compound ranking method	130
8.1	Mapping units	142
11.1	Example of functional classification	210
12.1	Types of mis-phasing	236
15.1	Deflection measurements using a benkelman beam	304
16.1	AGETIP agencies	311
16.2	Quality and cost-based consultant selection	318
17.1	Typical powers delegated to the resident engineer	326
17.2	Quality control activities	329
17.3	Summary of contractor's monthly statement	334
17.4	Example of determination of extension of time	338
17.5	Examples of claims for delay	340
18.1	Organization of equipment maintenance operations	365
19.1	Typical strategies for collecting condition data	377
19.2	Example of a treatment-based prioritization method	387
19.3	Summary of some prioritization methods	388
19.4	Problems with system implementation	391
21.1	Example application of HDM-4 in project analysis	431
21.2	Example application of HDM-4 to programme analysis	432
21.3	Example application of HDM-4 to strategy analysis	436
22.1	Clarification of budgeting concepts	444
22.2	Requirements for a successful road fund	447
22.3	Evolution of road administration	450
22.4	Sectoral roles	453
22.5	Commercialization of management practices	457
23.1	Problems with counterpart training	464
23.2	Training levels	466
23.3	Example of prioritizing training needs	470
23.4	Determining the form and content of a course	474
23.5	Checklist for course preparation	479
23.6	Communication	480
23.7	Common problems when instructing	481
24.1	Aid management in Danida	493
24.2	Identification	496
24.3	Preparation	497
24.4	Tools used to assist supervision of implementation	499

Contributors

Tonny Baek

Senior Technical Advisor
Danida
Ministry of Foreign Affairs
Asiatisk Plads 2, DK-1448
Copenhagen, Denmark

Chris Baguley

Road Safety Specialist
International Division, TRL
Crowthorne House, Nine Mile
Ride Wokingham, Berkshire
RG40 3GA
United Kingdom

Peter Broch

Transport Economist
Asian Development Bank
6 ADB Avenue, Mandaluyong
City 0401 Metro Manila,
Philippines

Kent Falck-Jensen

Chief Highway Engineer
Cowi
Parallelvej 2, DK-2800 Kongens
Lyngby, Denmark

Dr Tue Hald

Development Engineer
Vestas Wind Systems
Paludan Müllers Vej 82, DK-8200
Århus, Denmark
(formerly Senior Hydraulic Engineer,
Rambøll)

Jan Hassing

Senior Water Resource Planner
DHI Water & Environment
Agern Allé 11, DK-2970 Hørsholm
Denmark

Dr Michael Høgedal

Siting and Offshore Specialist
Vestas Wind Systems
Paludan Müllers Vej 82, DK-8200
Århus, Denmark
(formerly Head of Ports & Coastal
Engineering, Rambøll)

Dr Goff Jacobs

Senior Consultant
International Division, TRL
Crowthorne House, Nine Mile Ride
Wokingham, Berkshire RG40 3GA
United Kingdom

Alan Jacobsen

Senior Water Supply Planner
Cowi
Parallelvej 2, DK-2800 Kongens
Lyngby, Denmark

Dr William V. Kennedy

Director
Commission for Environmental
Cooperation
393 Rue St-Jacques Ouest
Montreal, Quebec H2Y 1N9
Canada
(formerly Head, Environmental Policy
and Strategy, European Bank for
Reconstruction and Development)

Professor Henry Kerali
The World Bank
1818 H Street NW, Washington DC
20433, USA
(formerly Head of Highways Group,
the University of Birmingham)

Dr Jan Kildebogaard
Programme Co-ordinator
Royal Danish Embassy
PO Box 11243, Kampala, Uganda

Dr Richard Koole
Portfolio Manager, Asphalt and
Pavement Design
Shell Global Solutions
Route Departmentale 3, PB 97, 76650
Petit Couronne, France

Dr Britha Mikkelsen
Senior Sociologist
Cowi
Parallelvej 2, DK-2800 Kongens
Lyngby, Denmark

Torben Mikkelsen
Senior Highway Engineer
Kampsax
Parallelvej 2, Dk-2800 Kongens
Lyngby, Denmark

Dr Knud Rask Overgaard
Senior Consultant
Granstuevej 25, DK-2840 Holte
Denmark
(formerly Hoff & Overgaard)

Jens Erik Bendix Rasmussen
Chief Technical Adviser
Danida
Ministry of Foreign Affairs
Asiatisk Plads 2, DK-1448
Copenhagen, Denmark

Dr Richard Robinson
Consultant and Honorary Senior
Research Fellow at the University of
Birmingham
40 Larkwood Drive, Crowthorne
Berkshire RG45 6RL
United Kingdom

Dr John Rolt
Chief Research Scientist
International Division, TRL
Crowthorne House, Nine Mile Ride
Wokingham, Berkshire RG40 3GA
United Kingdom

Professor Bent Thagesen
Ejlskovvej 5, DK-4000 Roskilde
Denmark
(formerly Technical University of
Denmark)

Jan de Veen
Senior Adviser, Employment and
Investment Policies
International Labour Office
Offenbachlaan 38, 4384 ME Vlissingen
The Netherlands

Tim Waage
Chief Resident Engineer
Flat A6, Marine Gate, Marine Drive
Brighton BN2 5TP, United Kingdom
(formerly Kampsax)

Preface

This book is a new edition of *Highway and Traffic Engineering in Developing Countries*, published in 1995. It aims to meet a clear need for a textbook on planning, design, construction and maintenance of roads in the traditional developing countries of Africa, Asia and Latin America. Most of these countries are situated in the tropics, where natural conditions are different from those in temperate regions. This impacts on issues such as soils, construction materials, hydrology and drainage, and on the road design and treatment methods that are appropriate. This new edition has extended the focus to include those countries of the former Eastern Bloc, which share many institutional issues and the financial problems confronting countries in the 'South'. Most existing textbooks on road engineering are biased towards the experience and needs of industrialized countries with temperate climates, or they deal with specific problems, such as road building in the tropics, soil stabilization or terrain evaluation.

The book is designed primarily as a fundamental text for civil engineering students, but an additional objective is to offer a broader view of the subject for practising engineers in road departments, consultancies and international organizations. It provides a comprehensive account of the wide range of technical problems confronting road engineers working in the third world. However, it recognizes that engineering in this part of the world needs to be undertaken within the wider framework of development, institutional and social issues. As such, additional chapters are provided on roads and development, policy, environment, institutional development, training and development aid.

All chapters in this new edition have been updated substantially, new material has been added and the book has been restructured. The chapter on 'Roads and development' has been extended to cover issues of poverty alleviation and transition to a market economy. Entirely new chapters have been added on 'Policy', including the setting of performance indicators, and on 'Roads and the environment'. The 'Traffic', 'Traffic safety' and 'Economic appraisal' chapters have been entirely rewritten. A new chapter has been written covering the use of the HDM-4 road development and management tool, which has been issued since the last edition of the book. The 'Asphalt pavement materials' and 'Structural design of asphalt pavement' chapters have been updated to reflect improvements in design methods since the last edition. The 'Contracts and works procurement' and 'Contract supervision' chapters have been updated to take account of the recent revisions to the FIDIC Conditions of contract. A section on equipment management has been added to the 'Appropriate

technology' chapter. The chapters on 'Maintenance management' and 'Institutional development' have been entirely rewritten to reflect the recent advances in each of these areas, with greater focus on road financing, organizational restructuring and the use of the private sector.

The book does not cover specifically work on bridges, footways, road furniture, etc. Neither does it purport to address issues of safety associated with the testing of road materials and construction and maintenance of roads. Readers are expected to establish appropriate health and safety practices, and determine applicability of national regulatory requirements prior to use of any method described in the book.

Acknowledgements

The text has been written with the assistance of a number of professionals with many years of experience gained in the roads sector around the world. Their names appear in the list of contributors, in the table of contents and under the headings of each chapter. All authors have contributed freely of their time and knowledge, and we are indebted to them all for their contributions.

In addition, Henrik Grooss Olesen and David Tighe (both of Cowi) contributed to the 'Roads and development' chapter and Thomas K. Thomsen (Carl Bro) contributed to the 'Planning methods' chapter. Dr Colin Oram (University of Warwick) contributed to the 'Appropriate technology' chapter and provided a photograph. Kristian Skak-Nielsen (BlomInfo) commented on the 'Soil investigation' chapter, John Hine commented on the 'Economic appraisal' chapter, Linda Parsley (TRL) commented on 'The HDM-4 road investment model' chapter, and the late Dr Tom Jones commented on several of the chapters in the construction section. R. Scott Hanna (Acres International) provided figures for inclusion in the 'Roads and the environment' chapter. Photographs were also provided by Annabel Davis and Dr John Rolt of TRL Ltd. We are most grateful for all of these inputs.

Many of the illustrations have been reproduced from other publications. The sources are quoted below the illustrations and at the end of each chapter. Every effort has been made to trace all copyright holders but, if anyone has been overlooked, the publishers will be pleased to make the necessary arrangements at the first opportunity.

Richard Robinson
Bent Thagesen



'Happy cyclists in Vietnam'. (Photo: Heine Pederson)

Part I

Planning



Chinese farmers on their way to market. (Photo: Bent Thagesen)

Roads and development

*Peter Broch, Britha Mikkelsen and
Richard Robinson*

1.1 The meaning of development

Roads may be constructed for many reasons. However, most roads are built to facilitate the transport of people and goods, and so to promote development. However, development is an ambiguous concept, and is defined differently by different people. Box 1.1 provides some background definitions.

Another set of important definitions, relating to economic and statistical concepts, is commonly employed in the description of development, and is described in Box 1.2.

Concepts

Gross national
income

Box 1.1 Development

Development theory

Development theory is a systematic conceptualization of the conditions that determine the change in third world societies. It is concerned with the process of development, rather than with the achievement of a particular state of development.

Development strategy

A set of prescriptions on how to initiate and implement a development process. The formulation of a development strategy must, necessarily, be based on a perception of the mechanisms that hinder or promote development; that is, on a development theory.

Development concept

An overarching idea about the desired direction of social change; that is, about the goals that the strategy is intended to promote. Examples of such goals are individual freedom (the dominant concept in liberalism) and equality (the dominant concept of socialism). Development concepts almost invariably are derived from either of the two dominant political ideologies of the twentieth century: 'capitalism' ('liberalism') or 'socialism' ('marxism/communism').

*Box 1.2 Economic and statistical concepts**Gross national income (GNI)*

This is the aggregate value of the balances of gross primary incomes for all sectors in a country in a given year.

Gross domestic product (GDP) and gross national product (GNP)

Prior to 2001, organizations used the term ‘gross national product’ – which is identical to GNI. The term ‘gross domestic product’ was also used. Both measured the combined income generated in a country in a given year. The difference between the two concepts was that GDP measured the total income generated from production inside the country, whereas GNP also included income received in the country but generated elsewhere, such as remittances from migrant workers. In many third world countries, this constitutes a major source of income and foreign currency.

GNI per capita (also known as per capita income)

This is the average income per person which, in practice, is related closely to a county’s level of development. International organizations, such as the UNDP, OECD and the World Bank, routinely classify the world’s economies into degrees of development according to their per capita income. Countries are grouped as: low, middle and high-income economies.

Country groupings

Low and middle-income countries form the bulk of the group of developing and emerging countries. However, a few of these countries, which are oil exporters, actually fall in the high-income group. In 2001, the boundaries between the three income groups, as defined by the World Bank, were:

- Low-income economies – 1999 GNI per capita of US\$755 or less
- Middle-income economies – 1999 GNI per capita of US\$756–9,265
- High-income economies – 1999 GNI per capita in excess of US\$9,266.

**Per capita
income**

Table 1.1 presents summary data on GNI per capita and population growth. The table shows that, during the 1990s, the poorest countries outside Europe and Central Asia experienced average economic growth rates at least as high as those in the high-income economies. This presents an encouraging development relative to the preceding 20 years. From 1970 to 1989, the general picture was one of stagnant, or even negative growth in the poorest countries, especially in South Asia and Sub-Saharan Africa. This was exacerbated by high and accelerating population growth. Yet, income differentials remain huge, with a factor of around fifty between the poorest and the richest economies.

**Income
inequality**

There are also large income gaps within individual third world countries. Hence, ‘underdeveloped’ does not mean that everybody in the country is poor. Rather, the implication is that a large proportion of the population is poor, whereas a minority

Table 1.1 Summary data on GNI and population growth by region

Country group	GNI per capita 1999 (US\$)	Annual growth of GNI per capita 1990–99 (%)	Annual growth of population 1990–99 (%)
South Asia	440	7.7	1.9
Sub-Saharan Africa	500	2.4	2.6
East Asia and the Pacific	1,000	7.4	1.3
Middle East and North Africa	2,060	3.0	2.2
Europe and Central Asia	2,150	–2.7	0.2
Latin America and the Caribbean	3,840	3.4	1.7
High-income economies	25,730	2.4	0.6

Based on data in: *World development report 2000/01* (World Bank 2000).

enjoy incomes and lifestyles that are comparable to those of the vast majority of the populace in developed economies. A large income gap within a country is a characteristic feature of underdevelopment. It can be argued that the world is underdeveloped, since only a minority of its population is affluent.

The GNI concept has been criticized for its interpretation of development in monetary terms. The critics maintain that GNI reflects the values of the industrialized societies, where the maximization of income and the accumulation of capital are considered the driving motives behind people's actions. In addition, direct comparison of per capita income between nations is not straightforward because GNI measures only those economic activities that involve exchanges in the market (i.e. are paid for). Hence, the value of activities such as subsistence agriculture and small holdings, which are the dominant occupations for a large proportion of third world populations, are not taken into account. Furthermore, prices differ widely between economies, especially when it comes to staple commodities that usually are much cheaper in the third world than in developed countries. As a result, the purchasing power of a dollar varies significantly between countries.

Such difficulties have led to the formulation of alternative measures of development. Important examples are the 'purchasing power parity' (PPP) method, which adjusts monetary values to take into account the varying purchasing power in economies, and the 'human development index' (HDI), which has been published by the United Nations Development Programme (UNDP) since 1990. The HDI is a weighted composite index that includes three essential components of human well-being: life expectancy, literacy and income expressed in PPP-adjusted US dollars. There is no precise link between per capita income and the HDI, although a broad correlation exists as is shown in Figure 1.1.

It may perhaps be more appropriate to define development as 'sustainable improvement in the quality of life'. Economic growth is undoubtedly a crucial aspect of development, but social redistribution and political reform may be equally important.

The focus and emphasis of development theories and strategies has changed over time, and it is helpful to review this in terms of evolutionary periods.

Relevance of
GNI

Alternative
measures of
development

Quality of life

Changes over
time

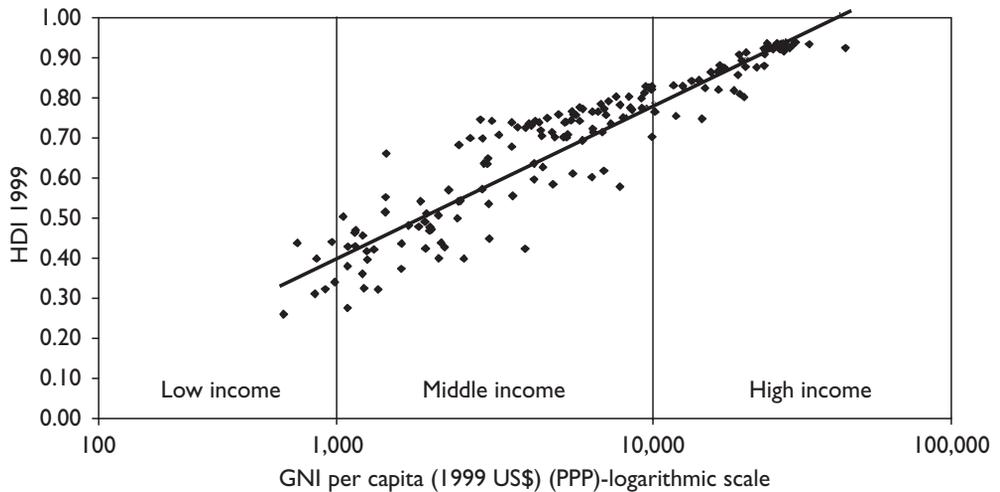


Figure 1.1 Relationship between GNI per capita and HDI.

Based on data in: *Human development report* (UNDP 2001).

1.2 Evolution of development theory

1.2.1 The postcolonial period: 1950s and 1960s

Modernization theories

Modernization theory was built on the concept that there were ‘traditional’ and ‘modern’ sectors of the economy in developing countries, operating independently. The concern was how to overcome traditional behaviour, values and attitudes, so that traditional social structures could be transformed into modern ones. An example of a modernization theory is that of ‘dualism’, initiated by the American economist Walther Rostow. He suggested that all societies have to pass through a number of stages before they reach a level of development comparable with western industrial societies. The time taken to reach a given stage depends on the society’s natural and economic conditions, although political and cultural phenomena do play a role.

Decolonization theories

An entirely different approach was taken by development thinkers from the third world itself, such as Gandhi (India) and Nkrumah (Ghana). The colonial history was seen as the major cause of underdevelopment, and the abolition of colonial linkages as a prerequisite for development. Important among the decolonization theories are the ‘dependency’ theories. These reject the idea that the modern and the traditional economic sectors function independently of each other and claim, to the contrary, that a network of exploiting relationship links the poorest peasant in the developing world with the big international companies. Thus, poor economies are an integral part of the capitalist economy. These relationships have been termed ‘economic imperialism’. The dependency theories had far-reaching implications for the anti-imperialist movement of the late 1960s and 1970s, and for the formulation of most national development strategies in those years.

Single-factor strategies

The early development theories typically identified a single factor as the main promoter for development. Nation building, the formation of a middle class, the

development of entrepreneurship, education, family planning and the provision of infrastructure, such as roads, are examples of factors that were singled out as being central to successful development. The most widespread explanatory factor was, however, that capital shortage was the primary barrier to development. This led to the formation of international financial institutions, such as the World Bank and the various regional banks. It has continued to dominate the thinking and actions of bilateral and multilateral donors since then.

In the 1960s, many countries embarked on a strategy of import substitution, with the aim of developing local industries to replace imports. An example is Nyerere's 'African socialism', where the governing principle was self-reliance. However, more often than not, such strategies failed to generate economic development because the resulting products were inferior to, yet more expensive than, comparable imports. One reason was that many countries lacked the skills required to build and operate complex industrial plants. Another reason was that domestic markets were small, thus preventing the attainment of economies of scale.

Import
substitution

Other countries, particularly those in South East Asia, where the population pressure is high and land is scarce, invested in scientific development of their agriculture. The process was initiated by the development and subsequent large-scale introduction of high-grade strains of rice. As a result, agricultural production in countries, such as the Philippines and India, increased significantly within a few years.

Green
revolution

Countries, such as Kenya, Singapore and South Korea, embarked upon liberal development strategies. These aimed to achieve integration into the world economy by encouraging the private sector to develop export-oriented industries and, in South Korea, included reform of the agricultural sector. In many ways, these strategies resembled the path of development of many industrialized nations during the nineteenth and early twentieth centuries. However, these liberal strategies also led to unequal social development and large-scale migration from rural to urban areas resulting in the formation of vast slum areas. The economies suffered repeated crises brought about by weak and inefficient civil institutions and a poorly developed banking system. However, in Taiwan, South Korea and Singapore, these strategies successfully elevated their economies to high-income status within a few decades, but success was not universal. Countries like the Philippines and Kenya got stuck at the lower end of the middle-income level.

Liberal
strategies

1.2.2 The 1970s and 1980s

The escalation of oil prices in 1972 resulted in increasing prices of manufactured imports, while lack of demand depressed prices for raw materials and agricultural produce, which were the main foreign currency earners in most developing countries. Foreign exchange revenues became insufficient to meet import requirements and service the often massive foreign debts, which many countries had accumulated while industrializing. Import substitution became unsustainable. The immediate response was increased regulation: price controls were imposed, imports and the access to foreign currency became tightly regulated and, in some countries, quotas were imposed on domestic distribution.

Collapse of
import
substitution

The debt crisis forced many governments to call on the World Bank and the IMF to gain relief and support with restructuring of their economies. The 'structural

Structural
adjustment

adjustment' programmes were inspired by the strategies adopted by the most successful countries: emphasis was placed on economic liberalization to increase the efficiency of local resources mobilization over that previously accomplished by state bureaucracies. The result was cutbacks in public services and subsidies, a reduction in regulation and abolition of price controls. A weighty criticism of the structural adjustment programmes is that they often lead to massive social degradation and escalating inequality.

Global
development
theories

The development theories of the 1980s emphasized global economic linkages as the driving force in development – in stark contrast to the earlier theories, which considered international economic links to be one-way channels for exploiting the third world. Examples of the new global theories concerned the 'internationalization of capital' and the new 'international division of labour'. These were both attempts to explain why industrialization in some developing countries had been successful, but had failed in others. These theories contain two dominant processes. On the one hand, diminishing rates of return in the industries of the developed countries triggered technological innovation, which enabled the transfer of complex production processes to developing countries. On the other hand, the governments of the newly industrialized countries (NICs) competed successfully for international investments by creating incentives for investors, such as access to cheap infrastructure in special export zones, tax holidays, and access to disciplined and cheap labour. The combined impact of these 'push-pull' processes has been to relocate labour-intensive industrial processes from the industrialized countries to the third world.

Inter-
dependence
strategies

The global perspective was also central to the Brandt Commission reports (named after the former German Chancellor) – *North-South: a Programme for Survival* was published in 1980, followed in 1983 by *Common Crisis North-South: Co-operation for World Recovery*. The reports recommended the transfer of massive resources from rich to poor countries to revive or spur their economic growth. The transfer would benefit both the receiving and the donor countries. The growing third-world economies would provide the developed economies with growing, external markets that, in turn, would allow the industrialized economies to prosper in the longer term. The perspectives of the Brandt reports were refined in the works of the South Commission: *The Challenge to the South*, published in 1990. This report maintains the global interdependency perspective, but shifts the analysis towards the internal weaknesses of financial mismanagement, institutional incapacity and weak human resources within the developing countries.

1.2.3 The 1990s

The end of
the cold war

The late 1980s saw the demise of the rivalry between communism and capitalism. During the cold war, both contestants granted largely unconditional support to friendly third-world governments in return for political allegiance. Since the purpose was to amass supporters, the receiving governments could be assured that they would receive ongoing support, even if their usage of aid was inefficient, and even if they paid only lip service to the grander principles of their respective sponsors. With the end of the cold war, most third-world governments on either side of the political divide soon saw their geo-politically motivated external support replaced by

strictly measured aid with strings attached, such as requirements for economic reform, human rights improvements and true social development. A wave of economic, political and legal reform ensued. This enabled the individual third-world governments to qualify for development aid and, more importantly, to allow them to attract foreign private-sector investments.

The transition of former communist countries towards a market economy has, in most cases, involved a clear outward focus through trade liberalization and openness to foreign investment. The resulting changes in the structure and direction of trade and the inflow of capital have been substantial, as has been the change in the structure of production. The impact of transition, however, is not confined simply to changes in output. Adoption of a market approach has dramatically increased the freedom of choice of consumers, producers, savers and employees. It has had a key influence on the role of the government, demonstrating the need to build strong institutions that support markets and provide social assistance to those losing out.

The 1990s represent a reversal to mainstream liberal economic theory as the basis for development thinking. The emphasis has been on establishing the conditions (good governance, free trade, deregulation, etc.) to allow the global and local markets to act efficiently, thereby creating growth and social development. There is also much emphasis on removing the many barriers that prevent the inclusion of the poor and the disadvantaged into the sphere of development.

The worldwide adoption of liberal economics, including the adoption of free market principles by most of the remaining communist countries, has largely closed the economic debate on development, at least for the time being. However, other concerns have arisen, particularly concerning the environment. These concerns are rooted in projections of emerging evidence that uninhibited economic growth may render our habitat, the earth, less hospitable and may, ultimately, threaten our very survival.

Other concerns have led to the emergence of globalization sceptics, in some cases developing into anti-globalization movements. Much of their thinking and vocabulary is inherited from the anti-imperialist movements of the 1960s and 1970s which, in turn, were inspired by the early anti-colonial thinking that developed in the former colonies in the 1950s and 1960s. This time, however, the opposition to globalization comes from the North. The sceptics are a highly heterogeneous group and their motivation for objection varies widely, from simple protectionism aimed at preventing the deterioration of ailing economic sectors in the industrialized countries, over objections to cultural imperialism, to more broad-based concerns about the unequal distribution of the world's resources and the general advancement of capitalism. Movements such as ATTAC have proposed comprehensive debt relief for poor countries and the subjection of international capital transactions to taxation – the so-called 'Tobin Tax'. The risk with such proposals is that the third world may be deprived of a genuine and much needed opportunity for export-driven economic development. Yet, with the leading industrialized economies – especially the United States, EU and Japan – maintaining huge subsidies for their agricultural sectors and stiff quota and tariff protection for ailing industries such as steel, textiles and garments, one may sincerely doubt if the rich world will indeed honour their part of the globalization deal.

Transition to a market economy

Reversal to liberal strategies

Environmental concerns

Globalization sceptics

1.2.4 Changes in the development discourse

Causality

The switch from emphasizing the ‘traditional’ segments of societies, to emphasizing ‘the poor’ has been an explicit change in the development discourse over time. This is not just simply a matter of semantics. In the first case, the causality was that people are poor because they are traditional, which indicates that the approach to development should be to change the minds of people and the structure of society, and prosperity would follow. Today, the view is reversed: people are traditional because they are poor. This implies that poverty must be eradicated to bring about development and the change from traditional to modern. Likewise, an outward perspective has replaced the introspection of early decolonialism. Most developing countries are now in search of an active role in the global economy, rather than turning their backs on the world. This reflects the experience of the handful of countries that successfully leapt the barrier to prosperity during the last 50 years. All followed a route towards economic liberalization and integration into the global economy.

Current development goals

Two main development goals are now being pursued. Poverty reduction has become the overarching goal for development. The goal is shared by most poor countries, many international development agencies and their partners. All member countries of the United Nations have agreed to work together to achieve the ‘Millennium Development Goals’:

- Halving the proportion of the world’s population living in extreme poverty by 2015.
- Improving access to basic services, such as health, education and water supply.
- Tackling the HIV/AIDS crisis in poor countries.

In addition, those countries that have recently rejected communism also face the challenge of making the transition from a centrally planned to a market economy. Adequate infrastructure services are crucial to the achievement of all these goals (DFID 2002).

1.3 Poverty

1.3.1 Understanding poverty

Extent of poverty

Poverty is not the cause, but the effect of development difficulties: employment, environment, ecology, energy, food security, immobility and many more concerns create poverty. The number of people in poverty has increased steadily to over one billion in 2002. Although poverty is increasing in urban areas, it remains predominantly a rural phenomenon.

Relative and absolute poverty

Different definitions of poverty attach different meanings to the concept. Poverty is often defined as deprivation in relation to a norm, or with incapacity in relation to a given living standard. This suggests that poverty is a *relative* concept. It is also argued that there is an *absolute* core of poverty. Starvation and hunger relate to this absolute notion, but so does the ability to avoid social shame, and the inability to bring up or educate children. The concept of poverty may change over time, but its

core concern is the inability to fulfil fundamental needs. In India, for example, poverty has been associated with ‘lack of income and assets, physical weakness, isolation, vulnerability and powerlessness’. ‘Quality of life’ and ‘well-being’ are alternative terms that offer a more optimistic prospect for ‘poverty reduction’.

In the past, poverty was often defined in terms of wealth, and ‘wealth ranking’ was used as a means to target the poorest of the poor. However, ‘wealth ranking’ implies a materialistic focus on assets, and is questioned for having a bias relating to Western values (‘Euro-centric’). Referring instead to ‘well-being’ encourages a re-orientation towards ‘quality of life’ which is more in line with global values. Being poor in material terms does not necessarily indicate absence of well-being, which is culture-specific and difficult to quantify. Note that *wealth* or *well-being rankings* relate to households. However, in communities where everybody is subject to considerable stress, such as with refugees, for example, ranking on a quantified basis seems questionable. Beyond a certain limit, attempting to identify variations in stress, malnourishment and misery becomes irrelevant.

Wealth and well-being ranking

The need to address these concerns has resulted in a broadening of the definition of poverty beyond the notions of inadequate private income or consumption towards a more comprehensive perspective: poverty is ‘absence of a secure and sustainable livelihood’ (Mikkelsen 1995). The importance is now stressed in terms of

Broadened definition

- incidence of poverty; that is, vulnerability;
- severity of poverty; that is, absolute, severe and relative levels of poverty;
- distribution of poverty between individuals by gender and age, and over time;
- poverty as a process, since people move in and out of poverty;
- small sample measurements of poverty among persons;
- profiles of poverty groups’ access to privately or publicly provided goods and services;
- assessments of the influence on poverty groups of ‘external’ factors, such as structural adjustment programmes, food prices, social and family networks, etc.

Conceptualization of poverty in rural areas tends to distinguish between ‘community poverty’ and ‘individual poverty’, as in Box 1.3. Urban poverty is more likely to be defined in terms of individual poverty than at the community level.

Community and individual poverty

1.3.2 Measuring poverty

Poverty is measured conventionally by the income or expenditure level that can sustain a bare minimum standard of living. The ‘poverty line’ sets the measure of absolute poverty, and an upper limit of US\$370 a year (US\$1 per day in 1985 purchasing power dollars) has been used as the ‘poverty line’ (World Bank 2000). People whose consumption falls below this level are considered to be ‘poor’. A lower poverty line of US\$275 is also used, and those below this level are considered to be ‘very poor’. However, the actual values used are country-specific.

Quantitative indicators of poverty

Other quantitative indicators of poverty are also used. UNDP have included a ‘human poverty index’ (HPI) within their HDI, described earlier. This attempts to shift the focus of poverty from income deprivation to capability deprivation and

Other quantitative indicators

*Box 1.3 Community and individual poverty**Community poverty*

- physical isolation – the poor often live in isolated communities, with poor quality roads or no roads whatsoever, and little external institutional presence; this is exacerbated by seasonal contexts where rainfall affects the quality of rural roads and/or leaves rivers impassable; isolation reduces access to consumption goods and services, physical and social infrastructure, and production inputs including financial, extension services, and commodity markets;
- access to safe water;
- quality of land;
- social capital – access to social networks at the household, inter-household, community and societal level is the single most important asset of the poor.

Individual poverty

- individual ‘ascribed’ attributes – gender is a key distinction, with widows, single mothers and, to a lesser extent, female heads of households being perceived as amongst the very poorest; poverty status is often linked to access to and control over resources;
- hunger and nutrition;
- access to productive land;
- access to productive assets;
- access to health and education.

Adapted from: Brocklesby and Holland 1998.

impaired human functioning. The Development Assistance Committee (DAC) of OECD is one of the key forums in which major bilateral donors work together. They have set a goal of reducing by one-half the proportion of people living in extreme poverty by the year 2015. In this context, poverty is defined by

- incidence of extreme poverty: population with income below US\$1 per day;
- poverty gap ratio: incidence times depth of poverty;
- inequality: poorest fifth’s share of national consumption;
- child malnutrition: prevalence of underweight among five-year olds.

Qualitative measures

Whilst it is recognized that the international poverty index of US\$1 per day is a useful tool to raise awareness and to generate political momentum, other indicators have different and complementary uses in the identification of poverty. The challenge is to

achieve a trade-off between measurability, which requires standardization, and local complexity. The following poverty factors are now normally taken into account:

- Income or consumption poverty
- Human (under) development
- Social exclusion
- Ill-being
- (Lack of) capability and functioning
- Vulnerability
- Livelihood unsustainability
- Lack of basic needs
- Relative deprivation.

Considerations of poverty by some agencies have more recently emphasized other factors (Maxwell 1999). These include social exclusion, which focuses on multiple deprivation (low income, poor housing, poor access to health and education), but also exclusion from democratic and legal systems, markets, welfare state provisions, and family and community. The aim of the new approach is to do away with pre-conceptions about the aspirations of rural people, and to develop an accurate and dynamic picture of poor people and their environment. This provides the basis for identifying the constraints to livelihood development and poverty reduction. Such constraints can lie at the local level or in the broader economic and policy environment.

Sustainable
livelihoods

At the heart of the so-called ‘livelihood’ approaches is an emphasis on people (Ashley and Carney 1999). Sustainable livelihood approaches work with people, supporting them to build upon their own strengths and realize their potential. At the same time, they acknowledge the effects of policies and institutions, external shocks and trends. The approaches recognize the complexity of rural livelihoods, and the multiple dimensions of rural poverty in terms of its cause and effects. The livelihood approaches see sustainable poverty reduction as achievable only if external support works with people in a way that is consistent with their existing living environment and with their ability to adapt.

People-
centred

1.4 Evolution of road development

Development theories have been discussed in general terms. The implications of these theories for the road sector are now discussed. First, the evolution of road sector development is considered.

Road sector

After independence, significant investments were made constructing new trunk roads. This focus was dictated by the then prevailing development strategy of rapid industrialization. During the 1960s, the International Development Association (the concessional loan organization within the World Bank Group) spent 30 per cent of its total investments on transport infrastructure, mostly trunk roads. The impact was profound, as exemplified by Kenya where the length of the network of paved roads was expanded from about 1,000km in 1960 to about 3,000km in 1970. However, most of these new roads were under-utilized by normal standards and did not generate economic returns. They are now under-maintained and provide low and declining levels of service, further diminishing anticipated benefits.

Early
investments

Focus on agriculture	In the 1970s and 1980s, the focus gradually turned towards the agricultural sector, and a larger proportion of road investments was directed towards rural access roads. The aim was to stimulate agricultural production and improve the conditions of rural life. Large-scale donor-funded integrated rural development programmes were enacted to improve the mobility of the rural population and stimulate farmers' integration into the market economy. Again, because of inadequate maintenance, roads constructed under these programmes often deteriorated rapidly.
Sustainability	During the 1980s, it was realized that countries had only a limited capacity to finance maintenance of their road infrastructure. Attention was switched towards the preservation and continued upkeep of existing roads rather than the provision of new facilities. The only roads taken care of were those deemed economically and institutionally sustainable in the long term. At the same time, the structural adjustment programmes supported this trend by refocusing public budgets away from investment and towards recurrent expenditure, including maintenance. By the end of the 1980s, a new type of road fund was introduced to raise funds for road maintenance. Revenue for these funds comes principally from a fuel levy. Road users are involved directly in the management of the fund through a 'roads board'. In return for the road users' acceptance of paying a levy, their elected representatives gain direct control of prioritizing road expenditure.
Traffic increases	The 1990s saw the re-invigoration of many third-world economies and the entry of countries from Eastern Europe into the market economy. In the middle-income countries, the lifting of import quotas, which for decades had kept domestic markets under-supplied with vehicles, led to rapidly increasing flows of traffic. Many major cities saw massive traffic congestion. At the same time, the (hesitant) reduction of the trade barriers in some of the rich economies led to improved agricultural prices and increased demand. Some governments have turned to the private sector to finance, build and operate toll roads to provide the required increase in traffic capacity. These have preserved scarce government and donor funds for other projects with little or no commercial attraction, such as roads in rural areas.

1.5 Impact of roads on economic development

Transport and economic growth	Transport investment reduces the cost of raw materials, labour and other products, reducing the cost of production directly. Improvements in transport extend the distance to break-even locations, thereby expanding the area of land under cultivation, and expanding the production of exports. Reduced cost and improved quality of services should also reduce the delivered price of products and, hence, promote regional and international trade. Resulting increases in farm-gate prices should raise farmer incomes, although the extent of this depends on the competitiveness of the transport sector market. Transport investment also contributes to economic diversification, and increases the economy's ability to handle risks (Gannon and Zhi Liu 1997).
Road density	However, the road networks of developing and emerging countries are sparse compared with those of developed countries, yet their upkeep and development require a much larger share of available resources, as shown in Table 1.2. The table does not take account of the differences in land area, population density, degree of

Table 1.2 Road densities for selected countries

Country	GNI per capita (current US\$)	Length of paved road (km per million population)	Length of paved road (km per GNI in billion current US\$)
Ethiopia	103	61	4,441
Burkina Faso	235	176	4,690
Tanzania	266	113	10,068
Bangladesh	360	150	4,377
Zimbabwe	471	730	3,270
Indonesia	683	776	2,455
Honduras	859	439	2,508
United Kingdom	24,231	6,245	258
Denmark	32,723	13,418	410
United States	33,087	13,416	690

Adapted from: World Bank (2000).

urbanization and many other relevant factors that characterize the individual countries. However, it does provide an indication of the relationship between per capita wealth and per capita road provision.

This relationship may be interpreted in two ways: the supply of paved roads promotes wealth, or wealth allows a nation to provide its population with superior roads. Both interpretations are valid, since the relationship between roads and development is not a simple one-way causality, but is a series of complex interactions over time. The budget constraint on road provision is illustrated in the last column of Table 1.2. This shows that roads put a much higher maintenance burden on poor countries than on rich ones, since the maintenance of each kilometre of road represents a much larger share of national income. Against this background, the perennial difficulties with securing sufficient funds for the maintenance of roads in the third world are hardly surprising.

Road works can be carried out with technology using various mixes of labour and machinery. 'Labour-based' methods use a high proportion of labour in undertaking works and, in appropriate situations, can offer benefits over methods using large amounts of equipment (see Chapter 18). The localized benefits of labour-based methods extend beyond the cost savings of road works and the creation of jobs. Other benefits include savings in foreign exchange, injection of cash into local economies and transfer of knowledge of road works to local communities.

Interpretation

Labour-based
methods

1.6 Transport and poverty

1.6.1 Basic issues

One of the first studies to recognize the impact of rural transport programmes on poverty was that of Howe and Richards (1984). This identified that wealth is highly

Immobility and
poverty

correlated with mobility, and poverty with immobility, in terms of access to both roads and vehicles. The conclusions drawn by Howe and Richards in 1984 are still pertinent today:

- Local circumstances have a considerable effect on the manner in which new road construction and improvement influence economic and social change; for this reason, it is not easy to predict the outcome of investment from experience gained elsewhere.
- The continuing optimism with which most road investment programmes are still regarded in relation to their effect on poverty cannot be sustained by the evidence.
- Changes that can be expected from road improvements are seldom as significant as those likely to result from making a trafficable route available for the first time.
- If contribution to agricultural output is the main criterion for rural road selection, then investments can be expected to reinforce stratification of existing social and economic structures; this is because investments help wealthier and better-informed producers to expand faster than others.
- Even where vehicle operating cost changes appear substantial, they are unlikely to lead to significant changes if crop development is controlled arbitrarily, rather than regulated by market forces, if transport services are subject to monopoly control, if the rules governing ownership and entry to commercial vehicle operation are unduly restrictive, or if transport comprises only a small proportion of the final cost of a product.

Transport development and poverty reduction

Transport can contribute to poverty reduction through its indirect impacts on economic growth, or its direct impact on personal welfare of the poor (Gannon and Zhi Liu 1997). Investment in the transport sector improves access to economic opportunities by reducing transport costs. Local access roads have a direct and significant impact on the daily life of the poor. The provision of transport services, including the construction and maintenance of transport infrastructure, generates demand for (often unskilled) labour and provides income-earning opportunities for the poor.

1.6.2 Access

Access and poverty

Poverty assessments from many countries point to the pivotal role of access as an enabling condition for development in all sectors. For rural dwellers, better access to employment opportunities, health centres, hospitals, schools, district offices and markets are often more important than measures aimed directly at economic development. A network of unclassified roads, tracks, trails and paths provides basic accessibility for the rural poor. Most rural transport infrastructure has an earth or gravel surface and serves motorized flows that are often exceeded by the flows of pedestrians, cycles and animals (Gwilliam and Shalizi 1996).

Mobility and service quality

Rural participants link physical isolation to limited mobility, inadequate geographical coverage of services and low levels of community outreach. Transport access and quality are key determinants of degrees of isolation in remote rural communities. Improving rural mobility and ease of transport should therefore be part of any poverty alleviation programme (Malmberg Calvo 1998).

Thirty-three per cent of China's population and 75 per cent of Ethiopia's still do not have access to all-weather transport. Walking more than 10km a day each way to farms, schools and clinics is not unusual in rural areas, particularly in Sub-Saharan Africa (Gwilliam and Shalizi 1996). Particularly in climatic zones with high seasonal rainfall, rural roads tend to deteriorate during the rainy season, limiting access to specialist health care, and compounding feelings of social and physical isolation as communities are effectively cut off (Brocklesby and Holland 1998). Research in Ghana has shown that providing reliable basic access, even on the simplest of roads, gives much greater economic returns than improving surfaces on low volume roads (Hine *et al.* 1983). However, increasing access for the rural poor requires more than the expansion of tertiary transport networks. It also requires the provision of more public transport services.

Enhancing
accessibility

1.6.3 Rural transport

An issue often overlooked, when planning roads, is that benefits will only be reaped if transport is forthcoming for both people and goods. Often projects identify the need for better transport to overcome social issues, but then recommend the need for road improvements (Howe and Richards 1984). Concern with vehicles per se rarely goes beyond the routine prediction of motorized traffic growth and generation. Decision-makers involved in the planning and financing of rural roads should view the movement needs of the poor as transport problems, relating to both vehicle and track, rather than simply as accessibility difficulties which can be solved by the provision of roads.

Vehicle issues

The poorest members of communities still walk as they cannot afford the available means of transport, so rural transport is largely non-motorized. Because the poorest groups are not able to afford any for-payment transport services, they have to carry heavy loads on their backs and heads over long distances. Typical loads include agricultural inputs and outputs, water and fuel for home needs. Non-motorized transport makes use of wheelbarrows, cycle carts and various forms of animal power. The dominant mode in China and some of the smaller poorer countries of South East Asia is the bicycle. Even in the public transport sector, the cycle is dominant in some countries. In rural Africa, non-motorized transport carries 90 per cent of the freight (Gwilliam and Shalizi 1996). Non-motorized transport is associated with poverty but, despite this, its users have been disadvantaged in various ways:

Non-
motorized
transport

- Approaches to road development tend to be top-down, concerned with the transport system as a whole, and addressing primarily the needs of motorized transport and the wealthy (Howe 1997).
- Government fiscal discrimination against non-motorized transport results in a lack of financing and manufacturing capability (Gwilliam and Shalizi 1996).
- Physically, vulnerability results from the failure to separate motorized and non-motorized traffic on the road (Gannon and Zhi Liu 1997).

One of the best ways to help the poor is thus to improve non-motorized transport. Switching from walking to cycling can make dramatic improvements. For passenger transport, cycling offers speeds of at least three times, and effective ranges of movement nine times as great as walking. For freight transport, a study in Ghana

Non-
motorized
transport
interventions

showed that it takes two person-days to move one tonne-km by head-loading compared with one person-hour with a bicycle and trailer (Hine *et al.* 1983). Improving non-motorized transport can only be achieved by a bottom-up approach to road development (Howe 1997).

1.6.4 Transport needs of women

The role of women

Particular problems arise in meeting the transport needs of women. In rural areas, most women are engaged in agricultural work and some small-scale commodity production and, particularly in Africa and the Caribbean, in marketing the produce of the family. They are responsible for collecting water and fuel. Studies in Mozambique and Tanzania show women spending more than four hours a day solely on local transport, which is three times the average for a man (Edmonds 1998). Women who are mothers make more trips than men for educational, health and other welfare purposes. Essential trips are also more dispersed in time and location than those of men. To date, however, transport policies have been geared primarily to the needs of men (Gwilliam and Shalizi 1996). Many of the trips made by women are in categories conventionally, and often incorrectly, regarded as inessential; that is trips not associated with formal work. They are disregarded in resource allocation and planning. Within the family, women have less access than men to bicycles or cars owned by the family, and are much more dependent on public transport and walking. In some countries, this is accentuated by strong taboos against women riding bicycles or driving cars (Malmberg Calvo 1998).

Eliminating gender bias

Reducing the transport burdens on rural women would release their time and energy for more productive and socially beneficial activities (Gannon and Zhi Liu 1997). Using the simplest form of wheeled vehicle would immediately halve the amount of time that women require for local transport. Making services more accessible in rural areas, either by improving the way that services are provided, or by increasing access to transport, benefits the whole family by releasing a woman's time. Failure to consider these possibilities for improving the lot of women often stems from inadequate analysis rather than excessive cost. Adopting a checklist of the effects that transport projects and policy reforms could have on the essential functions performed by women would make mainstream transport policy and planning more sensitive to these needs (Gwilliam and Shalizi 1996).

1.6.5 Sustainability

Sustainability as the basis of transport development

To be effective, transport development must satisfy three main requirements (Gwilliam and Shalizi 1996). First, it must ensure that a continuing capability exists to support an improved material standard of living. This corresponds to the concept of 'economic and financial sustainability'. Second, it must generate the greatest possible improvement in the general quality of life. This relates to the concept of 'environmental and ecological sustainability'. Third, the benefits that transport produce must be shared equitably by all sections of the community. This is termed 'social sustainability'.

Box 1.4 Selected recommendations in support of social sustainability

The following can facilitate increasing social sustainability of transport, by making poverty reduction an explicit part of national and local transport development strategies.

Target the transport problems of the poor

- Improve the physical access to jobs and amenities to reduce ‘excessive’ time spent walking.
- Reduce barriers to the informal supply of both passenger and freight transport, subject to reasonable and enforceable levels of safety.
- Enable greater use of non-motorized transport by improving infrastructure, and by eliminating fiscal and financial impediments to vehicle ownership.
- Eliminate gender bias by integrating the transport needs of women into the mainstream of transport policy and planning.

Improve the approach and criteria for addressing the transport problems of the rural poor

- Emphasize access, for example, by ensuring that bridges and culverts are durable and do not collapse, rather than high standards of performance, for example, by paving surfaces to increase speed.
- Support cost-effective labour-based methods for constructing and maintaining rural roads.
- Ensure community participation in decision-making on local transport investment and maintenance; establish extension services to provide necessary technical advice and training, and the support of rural development funds.

Protect the poor against the adverse effects of changes in general transport policies and programmes

- Minimize or, where unavoidable, mitigate the effects of resettlement by ensuring that people displaced by projects are resettled expeditiously and fairly.
- Mitigate the effects of any redundancy in over-staffed transport enterprises by ensuring that constructive re-employment and severance arrangements are in place.
- Develop efficient subsidy schemes for ‘social service’ public transport by defining public service obligations, and establish contractual compensation arrangements that are sustainable fiscally.

Adapted from: Gwilliam and Shalizi (1996).

Social sustainability

In rural areas, the poor are mainly dependent for their livelihood on their ability to produce and market agricultural products. Increased access to traded inputs (e.g. fertilizers and equipment) and the possibility of transporting agricultural products to distant markets create the conditions for cash-cropping to replace subsistence farming. This can also facilitate the development of non-agricultural activities in rural areas. This requires provision and maintenance of the infrastructure and services that are most critical in ensuring that the poor have access to markets, employment and social facilities. Some recommendations for social sustainability are given in Box 1.4.

A vision for transport

Howe (1997) considers that, if road development policy is really to address the needs of the rural poor, a new paradigm is needed. Road development should be part of an overall vision for transport that indicates how the supply of infrastructure and vehicles is to be provided, regulated and managed in relation to given demands. People, and their travel and transport needs, have to become the basis of any new transport development vision. It should be emphasized that this vision should extend to all the population, including pedestrians, cyclists and others who are not car owners or users. Effectively, the majority of the population is not in the motorized transport market. Rural populations need more affordable, available and appropriate transport than that currently available, even to enter the market.

1.7 Impact of roads on the transition to a market economy

Markets

Markets facilitate the exchange of goods between people. The existence of efficient markets is crucial for the functioning and development of a modern economy. Markets enable the exchange of goods and services, and enable supply and demand to be matched at fair prices – provided there are sufficient potential participants to the exchange. This is because markets introduce competition, providing customers with choices about how their needs are met, and compelling providers to become more efficient. The risk of collusion among providers diminishes as the number of participants increases. To achieve a high number of participants, widespread access and high mobility are needed. In turn, this requires comprehensive infrastructure and low costs of carriage. Without good access, the barriers to market entry are prohibitive and only few people can participate. Without low cost of carriage, many will not be ready to pay the cost of participation in the market, even if they do have physical access, because the profit that can be gained does not cover the cost.

The role of the road sector

In this context, the role of roads becomes crucial. They facilitate the execution of market transactions by enabling interested parties to meet and by enabling the subsequent transport of traded goods to their place of usage. Many sectors rely on roads for their effective operation. A good road transport system is essential to the effective development of a market and is, therefore, a pre-condition to sustainable development of an efficient market economy. A good road transport system must function as a network, linking smaller tracks and roads with

each other and with larger roads. This requires the development of effective finance, management and operation of roads. The road industry will be among the largest of any industry in a country. This size is masked because of its public sector nature, and because operations are financed from taxation through the government budget process. There are significant potential benefits in increasing transparency of expenditure, and in improving sector effectiveness and efficiency.

However, the provision of good infrastructure on its own is not sufficient. Good legal structures and dependable law enforcement are equally important. These often leave much to be desired in the third world. As a result, collusion, corruption, extortion and nepotism are rampant in many places, and detract from the efficiency of the market. Good infrastructure cannot on its own cure these problems but, without good infrastructure, misbehaviour thrives, shielded by the lack of cost-efficient access. There is a need for restructuring of inefficient government institutions through commercialization, and for introducing private sector operation and management in all cases where competition can be achieved. There is a need to strengthen management by introducing sound business practices and encouraging managerial accountability. This will have implications for the number of staff employed in the public sector because of increased use of the private sector, and increasing levels of competition for services (see Chapter 22).

Institutional
factors

References

- Ashley, C. and Carney, D. (1999) *Sustainable Livelihoods: Lessons from Early Experience*, London: Department for International Development.
- Brocklesby, M.A. and Holland, L. (1998) *Participatory Poverty Assessments and Public Services: Key Messages from the Poor*, London: Department for International Development.
- DFID (2002) *Making Connections: Infrastructure for Poverty Reduction*, London: Department for International Development.
- Edmonds, G. (1998) *Wasted Time: the Price of Poor Access*, Geneva: International Labour Office.
- Gannon, C.A. and Zhi Liu (1997) 'Poverty and transport', *Transportation Water and Urban Development Papers TWU-30*, Washington, DC: The World Bank.
- Gwilliam, K. and Shalizi, Z. (1996) 'Sustainable transport: priorities for policy reform', *Development in Practice*, Washington, DC: The World Bank.
- Hine, J.L., Riverson, J.D.N. and Kwakye, E.A. (1983) 'Accessibility, transport costs and food marketing in the Ashanti Region of Ghana', *TRRL Supplementary Report 809*, Crowthorne: Transport and Road Research Laboratory.
- Howe, J. (1997) *Transport for the Poor or Poor Transport?* Geneva: International Labour Office.
- Howe, J. and Richards, P. (eds) (1984) *Rural Roads and Poverty Alleviation*, London: Intermediate Technology.
- Malmberg Calvo, C. (1998) 'Options for managing and financing rural transport infrastructure' *World Bank Technical Paper No 411*, Washington, DC: The World Bank.
- Maxwell, S. (1999) 'The meaning and measurement of poverty' *ODI Poverty Briefing*, London: Overseas Development Institute.

Mikkelsen, B. (1995) *Methods for Development Work and Research: A Guide for Practitioners*, New Delhi: Sage.

UNDP (2001) *Human Development Report*. New York: Oxford University Press for the United Nations Development Programme.

World Bank (2000) *World Development Report 2000/2001: Attacking Poverty*, New York: Oxford University Press for the World Bank.

Policy

Richard Robinson

2.1 The nature of policy

Policy provides a framework within which decisions can be taken about all aspects of road network management. The policy framework can help to ensure that decisions made by different organizations and bodies are achieving a common overall aim, and are consistent with each other. Policy can be considered to include all of the

Decision-making context

- policy statements and documents published by public administrative bodies;
- relevant laws and statutes, especially their preambles and announcements;
- decision of the courts and regulatory bodies on important issues;
- procedures and manuals issued by relevant organizations concerned with the overall management of the road network.

At its simplest, policy may be considered as the *what* and *why*. This contrast with ‘plans’, which are the *how*, *who*, *when* and *where*. Plans provide the *strategy* by which policy is implemented (Howe 1996).

In many organizations and bodies, policy is a very nebulous entity: it is poorly defined and ethereal. A structured approach to policy formulation is therefore necessary to overcome this. A well-structured policy is particularly helpful when budgets are constrained, because ownership and responsibility are clarified, and objectives can be set that are transparent, equitable and which reflect road user requirements more accurately. Clear policy also provides a firm basis for planning, for considering options and priorities, for determining physical achievement and obtaining value for money. Policy can be a powerful vehicle to facilitate change. But policy formulation is complex, because of the need to co-ordinate between many sectors, organizations and bodies if sustainable implementation is to be obtained.

Structured policy

Policy for road management needs to be set at two main levels:

- Government level, where political aims are defined, and legislation and other instruments are adopted which force or encourage implementation of the policy.
- Road administration level, where plans and procedures need to be put into place to implement the policy laid down by government.

Framework for decision-making

2.2 Government policy

Transport and road policy	Government transport policy provides a ‘statement of intent’ about how it intends the sector to be managed, operated and developed. It provides a framework within which organizations in the transport sector and road sub-sector can make decisions. Aspects of road policy can be set by national, state and local government administrations. The policy will provide the basis for determining issues such as the distribution of budgets, about priorities, and all other functions that are the responsibility of the particular road administration. Road policy should be part of an overall vision for transport. This vision needs to extend to all the population, including pedestrians, cyclists and others who are not car owners or users (Howe 1996).
Policy areas and responsibilities	Policies in different sectors may interact. For example, transport sector policy may have influence on social and economic development policy, and be influenced by fiscal policy. To be really effective, there needs to be a consistency of policy between all levels of government, and this requires that the policy formulation process is well co-ordinated between the various public administrative bodies concerned. This is difficult to achieve in practice.
Transport policy model	Figure 2.1 is an example of a transport policy model. It shows the relationships between general objectives of the society, transport objectives and possible means, which may be employed to achieve the objectives. Note that the policy model includes ‘direct means’, which comprise government budgets, transport taxation and regulations. Indirect means include similar government instruments in other sectors that impinge on transport.
Examples	Examples of government policy in the transport sector are shown in Boxes 2.1 and 2.2. In the first, from New Zealand, it will be seen that the policy covers issues relating to safety, cost-effectiveness, socio-economic and environmental aspects. It also states that a ‘commercial’ approach will be used for operations. The second example in Box 2.2 emphasizes, amongst other things, socio-economic goals and issues such as level of service, asset preservation, costs, safety and the environment. It will be seen from the items included in the policy statements, and also by those items excluded, that government policy is providing a framework for operations in the sector.

Box 2.1 Extract from a government policy statement from New Zealand

The New Zealand transport system will operate

- in a safe manner, in regard to operators, users and the general public, and with appropriate regulatory intervention being undertaken in such a way that maximizes the benefit of safe operation at reasonable cost;
- in a commercial environment that encourages long-term cost-effective innovation, flexibility, customer service and replacement of assets;
- in a manner that recognizes the transport needs of all elements of society and makes appropriate long-term provision for these needs;
- in a manner that recognizes the need for long-term environmental sustainability in every aspect of its activities.

Adapted from: New Zealand Ministry of Transport: Transport directions 1994–99: strategic directions for the next five years.

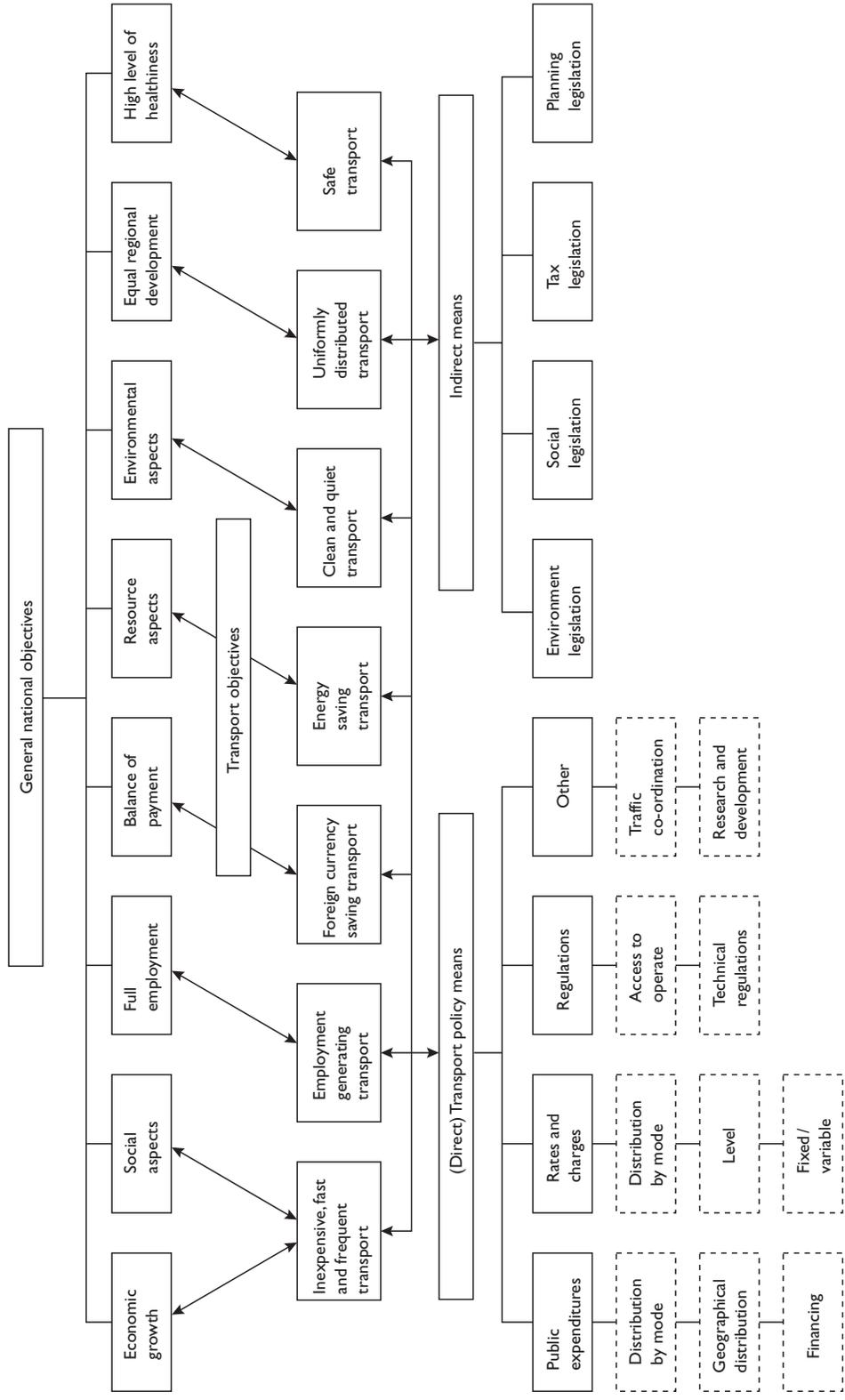


Figure 2.1 Transport policy model.
Source: Knud Rask Overgaard.

Box 2.2 Extract from a draft government policy statement from Zimbabwe

The Government of Zimbabwe will encourage the road sub-sector to manage its activities so as to sustain and develop an adequate road network in support of national and regional socio-economic goals by providing, maintaining, improving and regulating all roads in order to

- ensure access to centres of economic and social importance in rural and urban areas;
- ensure access to neighbouring states;
- provide an adequate level of service and minimize total transport costs;
- preserve assets vested in road infrastructure;
- provide an acceptable level of safety to road users;
- minimize detrimental impacts on the environment.

Adapted from: Government of Zimbabwe, Ministry of Transport and Energy: Road Sub-Sector Policy Green Paper, March 1999.

2.3 Organizational policy

Politics and administration

Government policy is implemented by public sector bodies operating within the sector, rather than by government itself. These organizations may be national, state or local government departments. Policy cannot normally be imposed on private sector bodies, other than through the use of policy instruments, such as legislation, regulation and taxation, or through contracts. Each organization operating within the sector needs to put in place its own policy framework to enable government policy, and its own policy, to be implemented. Whereas government policies provide a ‘political’ perspective, those produced by organizations need to be concerned with the administrative arrangements necessary to put policy in place within the organization’s own area of influence.

Structure of policy frameworks

Within an organization, a policy can be defined and formulated at a number of levels. Different approaches can use different numbers of levels, but the following framework is often adopted:

- Mission statement
- Objectives
- Standards.

Mission statement

The mission statement outlines, in broad terms, the nature of the operation being managed by the organization. It defines how one organization differs from another. The mission statement provides a summary definition of policy and how the organization aims to achieve this, so is normally kept fairly brief. It is mainly of interest to senior management staff within the organization, although it provides a basis of inspiration to all staff. To enable a mission statement to be put into effect, objectives and standards are required.

Each item in the mission statement should be supported by one or more objectives. These set specific goals to be achieved within the short- to medium-term (tactical) and long-term (strategic) time scales. Objectives are usually targeted at the managers in the organization who have the responsibility for delivering results in the form of work programmes. Often these will be the budget holders in the organization. Objectives need to meet the following criteria:

Objectives

- *Specific* – formulated in such a way as to be explicit, distinct and precise, in order to reduce the possibility of misinterpretation.
- *Measurable* – quantified in such a way that it is possible to determine whether or not the objective has been achieved.
- *Achievable* – such that it is actually possible for the organization to accomplish the requirement in the time available, or defined for response, with the resources available; amongst other things, this requirement means that there must be a match between objectives and the budget available to support them.
- *Relevant* – being pertinent to the organization’s mission, by having a direct bearing or influence on a particular item in the mission statement being considered.
- *Time-based* – achieved within a stated time scale.

Standards provide the detailed operational targets to be achieved by individual units in the organization. Standards may sometimes be supported by legislation or regulations. Standards are normally targeted at technicians, inspectors, supervisors, and others who have responsibility for ensuring that policy is implemented on the ground.

Standards

All standards should support an objective, and all objectives should be reflected in the mission statement. There should be a ‘one-to-many’ relationship between mission statement, objectives and standards, which thus provide a consistent set of criteria to guide all decisions. One particular reason for structuring policy in this way is because each component of the policy framework is applicable to a different customer, audience or vested interest. These customer interests, along with the relationship between the components of the policy framework, are illustrated in Figure 2.2.

Target audience for policy framework

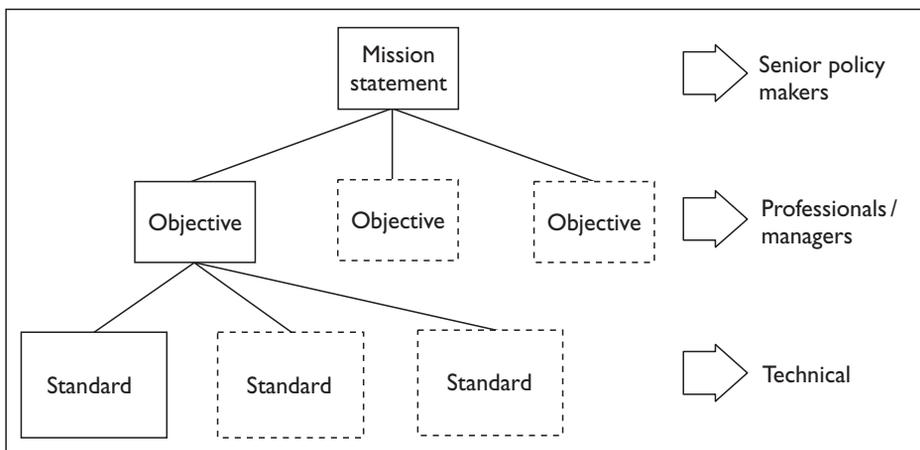


Figure 2.2 Relationship between components of the policy framework.

Source: Robinson et al. (1998).

Operation of policy

If the policy framework is well defined, the role of professionals in the organization becomes one of providing the appropriate technical solutions to implement the defined policy. Government then needs only to check that policy has been implemented by the professional staff, and to receive feedback from them about which areas of the sector policy need to be modified so that policy can be improved in the future.

2.4 Integrated policy

Relationship between levels of administration

The way that policy at different levels of public administration relate to each other is illustrated conceptually in Figure 2.3. This shows how a policy would be produced at a political level in central government, and by the various national bodies that would need to implement the policy. These might be a national ministry of transport, a national road administration, an environment ministry, for example. The policies

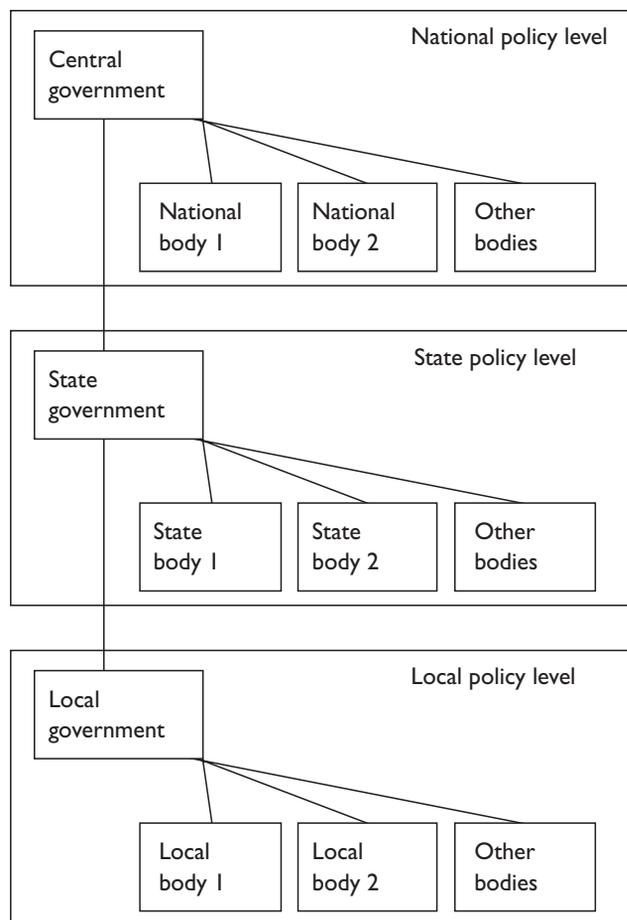


Figure 2.3 Example of relationship between policies produced at different levels of government.

produced by these bodies need to be consistent with and must amplify the policy produced by central government. This structure is mirrored at state and local government levels.

The relationships between the policy areas are complex, and it is often difficult to make a clear-cut distinction between where one policy ends and another starts. In an ideal situation, all of the policies in different areas and sectors, and at different levels of public administration, would be well co-ordinated and consistent. This is seldom the case because the different aspects of policy will reflect the different aspirations of policy formulators, with a wide variety of vested interests. For example, local policy may be at variance with national policy because needs are perceived as being different when viewed from local and national levels.

Complexity

An example of an integrated policy is shown in the policy extract in Figure 2.4. This shows how a national policy statement is supported by the policy framework of the road administration. The *national policy statement* provides a broad inspirational aim for the country as a whole. Each sector will need to support this through the development of its own policy. In this particular case, the roads sector is managed by the Public Works Department. Its *mission statement* indicates how it intends to operate and manage the sector such that national aims are met. This is a broad statement of intent that differentiates the requirements of roads from other sectors. It contains specific statements in the areas of level of service, safety, road user costs and road administration costs, each of which are supported by a set of *objectives*. The level-of-service objectives are illustrated. Where appropriate, each objective may need to be supported by *standards*. In the figure, the standards supporting the first two levels of service objectives are shown.

Example

It will be seen in the example that the different levels of the policy framework are targeted at different people. The national policy statement is targeted at the population in the country as a whole. The mission statement is targeted at the political level in government. The organizational objectives are targeted at those in the organization that are responsible for delivering the work programme. The objectives state the specific targets that are to be achieved. They are measurable, relevant to an element of the mission statement, specific and difficult to misinterpret. It can also be assumed that they are achievable within the budget and human resource constraints of the organization. Standards define the level of detail required by technical staff responsible, in the case of the example, for referencing the network, carrying out inspections of condition to determine need and in undertaking the work.

Audience for policy

Note that objectives and standards imply the need for a certain level of funding if the policy is to be *achievable*. A well-structured policy can be meaningful only if it is formulated in conjunction with a realistic assessment of the likely availability of resources.

Achievability

Each level of policy provides a framework within which people at the various levels can make the decisions necessary to carry out their activities. Provided the policy framework is adhered to, then decisions throughout the organization will be coherent and consistent and, ultimately, will be taken in a manner that enables national policy aims to be met. It is in this way that policy provides a tool for integrating decision-making.

Integrated decision-making

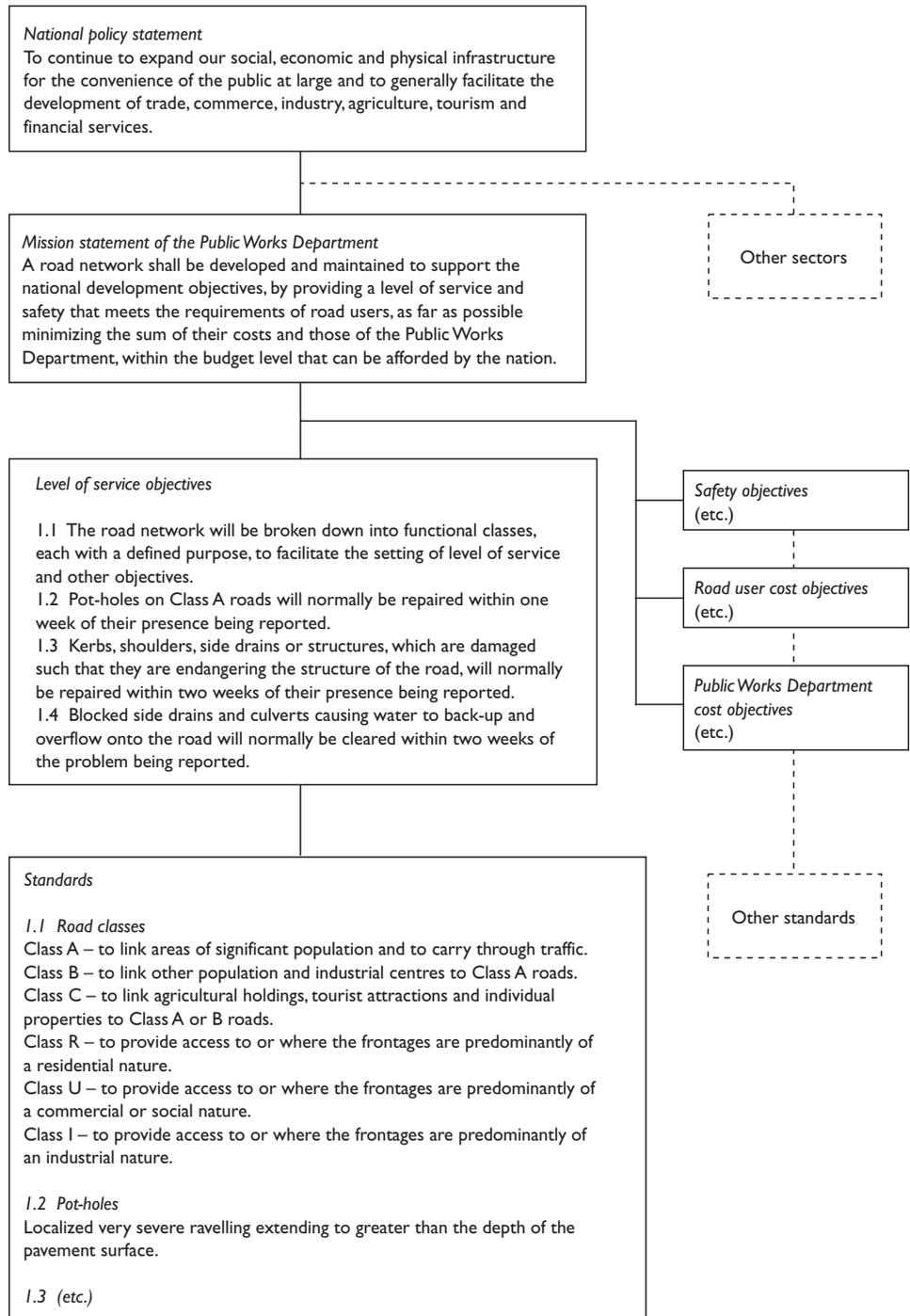


Figure 2.4 Example of national policy statement and organizational policy framework.

Adapted from: St Kitts Ministry of Communications Works and Public Utilities: Roads Code of Practice, 1995.

2.5 Dissemination

Policy only has meaning to the extent that it is implemented in practice. Thus, to be effective, the policy needs to be published so that it is available for public scrutiny. In some countries, policy is published in ‘white papers’, and is preceded by a ‘green paper’ stage where comments are sought through public consultation. Thus, there is a need to distinguish between the following:

- The policy itself – the aims and intentions of government in the area concerned.
- The means of disseminating the policy – policy statements, policy documents, and the like.
- The means or instruments for implementing the policy – laws, statutes, regulations and the like.

2.6 Performance indicators

2.6.1 Policy monitoring

Performance indicators can be used to assess the effectiveness of policy because they measure impacts, outcomes, outputs and inputs relating to decisions (OECD 1997). They help to quantify the degree to which objectives have been achieved, and to identify any constraints that are impeding the achievement of objectives. They help to answer questions such as (Talvitie 1999):

- Is the road administration doing the right things?
- Is the road administration doing things right?
- What things done by others significantly affect the road sub-sector?

Performance indicators differ from the operational statistics that most organizations produce, as shown in Box 2.3. The external focus of performance indicators means that they can normally be based on selected key objectives from the organization’s

Nature of performance indicators

Relationship to objectives

Box 2.3 Performance indicators and operational statistics

Operational statistics

These are used by the organization to oversee and manage all aspects of its operations. The statistics are mainly for internal use although, in the case of government departments, some may be made available publicly through their inclusion in annual statistical abstracts published by government.

Performance indicators

These relate to the impacts and perceptions by the ‘customers’ of the organization. As such, they are targeted externally. Indicators consist typically of a few selected operational statistics that characterize performance in a way that customers can relate to and understand.

Box 2.4 Examples of performance monitoring against policy objectives*Strategic planning*

Incorporating performance measurement into the design of any programme of activity forces greater consideration of critical assumptions that underlie the relationships and causal paths in the programme. Thus, performance indicators help to clarify the objectives and the logic of the programme.

Performance accounting

Performance indicators can help to inform resource allocation decisions if they are used to direct resources to the most successful activities and, thereby, promote the most efficient use of resources. They provide an incentive for an organization to increase effectiveness and efficiency.

Forecasting and early warning

Measuring progress against indicators may point towards future performance, providing feedback that can be used for planning, identifying areas that need improvement, and suggesting what can be done. For example, performance indicators can give early warning of undesirable trends and potential future problems.

Benchmarking

Performance indicators provide data against which to measure other projects or programmes. They enable comparative studies to be made within the sub-sector, between sub-sectors and sectors, geographic regions or other countries. They also provide a way to improve programmes by learning from success, identifying good performers and learning from their experience to improve the performance of others.

Synthesized from: Humplick and Paterson 1994; Cook 1995; Mosse *et al.* 1996.

policy framework. Thus, performance indicators present the key criteria against which the organization should be evaluated.

Examples

Some examples of how performance indicators can be used for monitoring against policy objectives are illustrated in Box 2.4.

2.6.2 Presentation of performance indicators

Since performance indicators are designed for a non-technical audience, the format of their presentation needs to be appropriate. An example of the type of performance indicators that can be produced is given in Box 2.5. This particular example is extracted from indicators published in a local newspaper, and indicates one method by which performance indicators can be disseminated to the customers and the travelling public. Some other formats for presentation are illustrated in Figures 2.5 and 2.6.

Box 2.5 Extracts from performance indicators published by a road administration

1	Road lighting	
	The percentage of street lights not working as planned	1.2%
2	Maintaining roads and pavements	
	The percentage of pot-holes repaired within 24 h	83%
3	Pedestrian crossings	
	The percentage of pedestrian crossings with facilities for disabled people	49%

Adapted from: Luton & Dunstable Herald & Post, 28 December 1995.

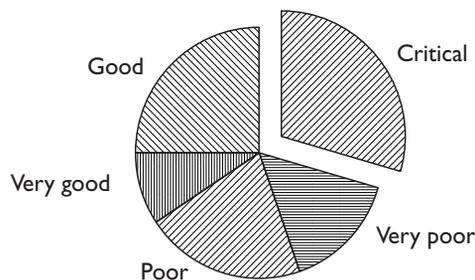


Figure 2.5 Example of a pie chart presentation of a performance indicator showing proportions of road network in different conditions.

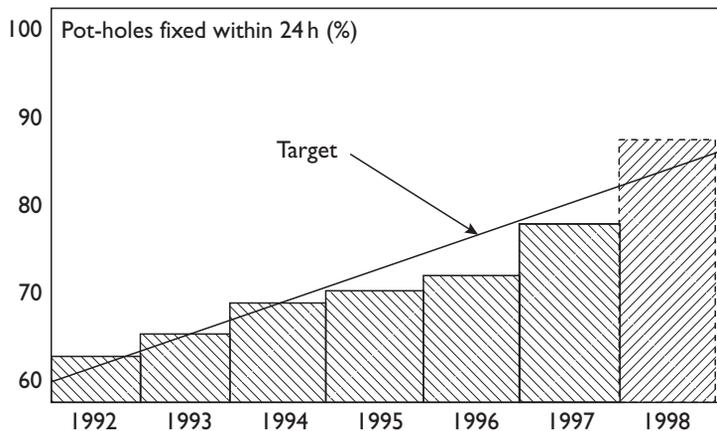


Figure 2.6 Example of a histogram presentation of a performance indicator showing achievement of pot-hole patching over time.

2.7 Policy formulation

Introduction of a new policy can have a significant impact on the existing administration and other vested interests. For policy reforms to have any chance of success, there must be a combination of political and public pressure for reform. Thus, policy

Commitment
and leadership

formulation requires strong leadership and commitment from the most senior levels of the administration, and this will need to be backed up further by support from government. The process can be facilitated if there is a ‘champion’ for the adoption of policy measures. A champion is often an individual holding a senior position; he or she will be a highly respected person and, preferably, will have the charisma to command support for the policy formulation process, and so be able to drive the process through to completion. However, it is also possible for a ‘champion’ to be represented by a group of committed people or an organization. Awareness needs to be created within potential interest groups and with the general public at all levels of society. Champions may need to find ways of fostering these interest groups and raising wider awareness of needs.

Steps in the approach

The formulation of policy needs to recognize that the large number of issues involved can make policy formulation a complex process. Because of the complexity of the various interacting factors, it may not be possible to produce an ‘ideal’

Box 2.6 Case study of policy formulation

Development of the road sector Green Paper in Zimbabwe provides a good example of the policy formulation process. This started in August 1996, when the heads of state of most countries in Southern Africa signed the *Protocol on Transport, Communications and Meteorology in the SADC Region*. This committed governments to making reforms in a number of areas, including road administration and finance. This was followed up in Zimbabwe with a workshop two months later, which involved stakeholders both from government and the road-user community. The workshop agreed to the outlines of a policy for reform. It set up a Steering Committee to guide the reform process, and sought donor funds for its implementation, including a Secretariat and consultancy contracts.

In March 1998, the Steering Committee held a workshop to draft a national policy statement. The outcome provided the basis of a Cabinet memorandum, putting forward the principles for a Road Fund Bill to Parliament. The main activities now centred on creating an awareness of the proposed reforms amongst stakeholders and the public. The Secretariat made visits to bodies throughout the country. Workshops were held in all regions. These were facilitated by the Secretariat, introduced by the Deputy Minister of Transport, and attended by a well-briefed media.

The process culminated in October 1998, when a further national workshop was held to finalize the policy statement. This was inaugurated by the Minister, and again attended by stakeholders from both public and private sectors. The conclusions were incorporated into a *Road Sub-Sector Green Paper*, issued in March 1999. This consultative document was widely publicized to seek feedback on the policy proposals. The aim was to finalize this as a Government White Paper. Legislation to enact the reforms would then be submitted to Parliament.

policy that meets all needs. A pragmatic approach must be taken: policies need to be put in place so that they address important needs in a rational way. Structured policy obviously benefits from a structured approach to its formulation, and there are two main steps involved. The first is the preparation of a national policy statement. The second is the preparation of policy frameworks that are consistent with this by organizations involved in the sector, at both national and local level. The policy formulation process needs to be managed proactively throughout the sector to enable a result to be produced that is coherent and integrated. The approach adopted for policy formulation can be similar to that for managing any institutional change (Talvitie 1997).

A key process for policy development is the holding of stakeholder workshops. These involve groups of individuals coming together to agree key issues and the policy measures that might be appropriate for addressing them. Although the aim of these is to produce policy documents, significant benefits result from the policy preparation process itself. Policy formulation is a time consuming process. There is a need for wide ranging discussions on the many issues affecting the road sector and for reaching a consensus on how best to address them. The types of issues will differ depending on the stage of the policy formulation process, and the level of detail discussed will increase as the process progresses. Workshops need to be ‘managed’ through skilled facilitation if the best is to be obtained from them. An example of how one country developed a structured policy for roads is given in Box 2.6.

Policy also reflects the changing needs and aspirations of society. As such, policy should be reviewed and updated on a regular basis. The impact of policy also needs to be monitored continuously to provide feedback on its effectiveness. Aims can be redefined to reflect actual achievements and lessons learned from experience. This may mean that different policies are set or updated at different times, so mismatches occur as needs are perceived to change in different areas. Thus, policy formulation is an ongoing process which allows incremental improvements to be made over time to reflect changing policy environments. Some key issues relating to the policy development process are illustrated in Box 2.7.

Stakeholder
workshops

Development
over time

Box 2.7 Key issues for policy and its development

- Policy, when developed in a formalized and structured way, can provide a framework within which effective and sustainable reform can develop.
- A policy-driven consultative approach to reform can generate ownership and commitment among stakeholders.
- Agreement of policy principles is easier than getting embroiled in discussing the details of legislation.
- Forming policy, and implementing change, take time – even to reach the *Green Paper* stage in Zimbabwe took over two-and-a-half years.
- The process requires commitment from the highest levels of government and needs a champion.
- Financial resources and full-time staff are needed to manage the process.
- Consultancy assistance can provide specialist support and facilitation.

2.8 Policy issues

The issues to be addressed by road sector policy will differ from country to country, and for different classes of road within the network. Some examples of the types of issues that need to be considered are given in Box 2.8.

Box 2.8 Examples of road sector policy issues

General issues for national-level policy

- Identification of the role of roads within the transport sector and, in turn, within the national economy; roles of roads of different classes.
- Access and mobility – requirements for level of service; measures to increase people’s ability to reach particular facilities, or bring facilities closer to rural communities; role of international transport links.
- Sustainability
 - ‘economic sustainability’ – measures aimed at preservation of assets vested in road infrastructure; requirements for investments to be subject to cost–benefit analysis; priority for funding to be given to maintenance; designation/adoption only of those parts of the network for which there is adequate maintenance funding; disposal of loss-making parastatals;
 - ‘social sustainability’ – measures to ensure an acceptable level of safety; emphasis to pedestrian travel; reducing the physical burden of transport on individuals; reducing transport requirements for women;
 - ‘environmental sustainability’ – minimization of adverse environmental impact; introduction of environmental protection legislation.

Infrastructure provision and management

- Responsibility for legal framework – identification of the road authority (i.e. the statutory representative of the state as legal owner of roads); requirements for designation/adoption of roads; minimum asset inventory requirements (gazetteer); measures to encourage community ownership/participation.
- Stakeholder representation in provision and management.
- Responsibility for sectoral organization – arrangements for administration, management and operation of the network; appropriate decentralized responsibility.
- Financing – means of financing major roads; securing budget transfers from central government for local roads; appropriate use of road funds; measures to encourage community financing and cost-sharing; the role for private finance; audit measures.
- Management, planning systems and local capacity – separation of government and administrative functions, and ‘client’ and ‘supplier’ functions within organizations; measures to encourage commercialization; use of competitive procurement.

- Technology – adoption of appropriate and cost-effective road standards; measures to encourage appropriate technology.

Means of transport

- Transport services – remove unnecessary regulatory constraints on vehicle construction and use, on routes and on operations; eliminate unnecessary import controls; introduce measures to support innovative transport service operations.
- Personal transport – measures to encourage the use of intermediate forms of transport.

References

- Cook, T.J. (1995) 'Performance measurement: lessons learned for development management', *World Development*, **23**(8), 1303–1315.
- Howe, J. (1996) 'Transport for the poor or poor transport? A general review of rural transport policy with emphasis on low-income areas', *IHE Working Paper IP-2*, Delft: International Institute for Infrastructural, Hydraulic and Environmental Engineering.
- Humplick, F. and Paterson, W.D.O. (1994) 'A framework of performance indicators for managing road infrastructure and pavements', in TRB, *Third International Conference on Managing Pavements, San Antonio Texas, 22–26 May 1994, Proceedings Volume 1*, Washington, DC: National Academy Press, 123–133.
- Mosse, R., Sontheimer, L.E. and Rogers, P. (1996) 'Performance monitoring indicators: a handbook for task managers', *Operations Policy Department*, Washington, DC: The World Bank.
- OECD (1997) *Road Maintenance and Rehabilitation: Funding and Allocation Strategies, and Performance Measures for the Road Sector*, Paris: Organisation for Economic Co-operation and Development.
- Robinson, R., Danielson, U. and Snaith, M. (1998) *Road Maintenance Management: Concepts and Systems*, Basingstoke: Macmillan.
- Talvitie, A. (1997) 'International experiences in restructuring the road sector', *Transportation Research Record No 1588*, Washington, DC: National Academy Press, 99–116.
- Talvitie, A. (1999) 'Performance indicators for the road sector', *Transportation*, **26**(1)Feb, 5–30.

Traffic

*Kent Falck-Jensen, Jan Kildebogaard and
Richard Robinson*

3.1 Introduction

Need for
traffic
information

Estimates of traffic flow are needed for most aspects of road planning, design and management. Traffic is a measure of the use of the road, and therefore provides important information about the needs of the road user or ‘customer’. The level of traffic will, for example, influence the road standards of geometric and pavement design, and the maintenance standard in terms of frequency of maintenance activities. Traffic data are also needed to forecast road performance. Various types of traffic information are needed for this, including

- Traffic composition
- Traffic flows and growths
- Axle loading.

Speed–flow
and capacity

As traffic flows increase, speeds will reduce and congestion will occur as the ‘capacity’ of the road is approached because of constraints of road geometry and other factors. The ‘speed–flow’ relationship will change depending on the time of day and period in the year. Determination of speed–flow relationships are needed when designing or improving high-flow rural roads, or roads in urban areas.

Accidents

Traffic accidents are also a form of traffic information, and are discussed in Chapter 4.

3.2 Traffic composition

3.2.1 Transport modes

Motorized
traffic

Traffic is characterized by a variety of transport modes. For the purposes of design and the evaluation of benefits, the volume of current traffic needs to be classified in terms of vehicle type. An example of a typical classification for motorized traffic is shown in Table 3.1. The distribution of motor vehicles in some selected countries is shown in Table 3.2.

Non-
motorized
traffic

In most developing countries, in addition to conventional motorized transport, a wide variety of non-motorized transport exists, including pedestrians and handcarts (Figure 3.1), bicycles and rickshaws. In addition, other categories of road users, such as cattle (single or in herds) may occupy road space, or even block the road completely.

Table 3.1 Example of a traffic classification

Class	Description
Cars	• Passenger vehicles seating not more than five persons, station wagons and taxis
Light commercial vehicles	• All two-axled vehicles with single rear tyres, not included as <i>Cars</i> or <i>Minibuses</i>
Minibuses	• Minibuses with 9–15 seats
Buses	• Purpose-built bus with more than 15 seats
Light trucks	• Two-axle truck, petrol driven, and with twin rear tyres
Medium trucks	• Two-axle diesel truck with twin rear tyres
Heavy trucks	• Three-axle diesel truck with twin rear tyres
Articulated vehicles	• Multi-axle articulated tractor and trailer

Table 3.2 The distribution of motor vehicles in some selected countries

	Number (1,000)				Network length (km)	Population (1,000)
	Cars	Buses	Commercial vehicles	Motorcycles		
Egypt	1,354	38	397	418	64,000	60,390
Kenya	278	19	62	32	63,800	25,480
Indonesia	2,409	595	1,391	10,091	336,377	200,500
Nigeria	885	903	912	441	193,198	103,175
Papua New Guinea	31	—	85	—	19,600	4,400
Philippines	703	1,130	220	882	161,264	69,952
Romania	2,408	46	365	329	153,358	22,770
Thailand	1,661	1,721	2,853	10,239	64,600	61,000
Denmark	1,797	14	282	52	71,321	5,251
USA	129,728	695	74,722	3,872	6,308,086	266,339

Note
Data from 1996.

Thus, the means of transport may differ considerably from that in industrialized countries. A wide range of low-cost vehicles has been developed to meet transport demand, particularly in the rural areas. Table 3.3 shows some basic vehicle types and their respective performance characteristics in terms of loading capacity, speed, range and relative cost.

3.2.2 Public transport

Public transport in developing countries is often a mix between well-organized bus companies (public or parastatal) and private minibus services – more or less well-organized. The bus companies usually operate a network of fixed routes with regular timetables and fares. The minibuses may be owned and run privately and individually with a high degree of flexibility. They may operate like taxis, competing for passengers in the free market. The advantage is that they can respond easily

Organization



Figure 3.1 Non-motorized transport. Photograph courtesy of TRL Ltd.

Table 3.3 Basic vehicle types and their performance characteristics

Vehicle	Load (kg)	Speed (km/h)	Range (km)	Relative cost
Carrying pole 'Chee-kee':	35	3-5	10	—
Traditional	70	3-5	10	10
Improved	70	4-5	10	20
Western wheelbarrow	120	3-5	1	20
Chinese wheelbarrow	180	3-5	3-5	20
Handcart	180	3-5	3-5	30
Bicycle	80	10-15	40	60-100
Bicycle and trailer	150	10-15	40	90-150
Bicycle and sidecar	150	10-15	40	90-150
Tricycle	150-200	10-15	40	150-200
Pack animal	70-150	3-5	20	Varies with specifications
Animal-drawn sledge (buffalo)	70-150	3-5	20	Varies with specifications
Animal-drawn cart (oxen)	1,000-3,000	3-5	50	100-180
Motorized bicycle	100-150	20-30	50	150-200
Motorcycle: 125 cc	150-200	30-60	100	250-600
Motorcycle and sidecar	250-400	30-60	100	350-800
Motorcycle and trailer	200-300	30-60	100	350-800
Motor-tricycle: 125 cc	200-300	30-60	100	500-1,000
Single-axle agricultural tractor and trailer	1,200	10-15	50	1,500
Asian utility vehicle	1,500	50-80	400	4,000

Based on data from: Habitat (1984).

Table 3.4 Public transport distribution in selected cities

City	Percentage of trips made by			
	Rickshaw	Minibus	Bus	Metro
Bangkok, Thailand	8	11	81	—
Calcutta, India	8	4	48	40
Chiang Mai, Thailand	7	86	7	—
Delhi, India	17	—	83	—
Jaipur, India	72	10	18	—
Jakarta, Indonesia	20	—	80	—
Kanpur, India	88	5	7	—
Manila, Philippines	—	64	36	—
Surabaya, Indonesia	54	39	7	—

Based on data from: Habitat (1984).

to the actual demand. Their disadvantage is that drivers and vehicles are difficult to regulate and control. This can often result in an inefficient use of road space and congestion in urban areas. The financial circumstances of operators often lead to aggressive driver behaviour and a disregard for safety.

Because of the low proportion of private cars, public transport often constitutes the backbone of the transportation system in both cities and rural areas. In Table 3.4, the distribution of trips among various means of public transport is shown in greater detail for some selected cities.

Distribution

3.2.3 Pedestrians and cyclists

A feature of traffic characteristic in developing countries is the large numbers of pedestrians and cyclists. As an example, bicycles account for 48 per cent of the total number of trips in Beijing (Zhao 1987). This clearly influences the capacity of the roads as well as the flow of traffic.

Volume

In most industrialized countries, bicycles are used mainly for individual trips and recreational purposes. In developing countries, very few people can afford a motor vehicle, and therefore the bicycle serves as a very important means of transport, not only for personal transport but also for the movement of goods. In addition, the bicycle has been developed for public transport in many cities. The Asian 'pedicab' is an example. The characteristics of some of the bicycle types are shown in Table 3.3.

Use

The value of a bicycle may equate to several months' salary for a skilled worker and, consequently, may be out of the reach of a large part of the population. For many people, walking is the only means of transport.

Walking

3.2.4 Application of design models

It is common for road design and traffic control measures to be geared to meeting the demand of motorized traffic in interaction with a limited number of pedestrians and cyclists. Large amounts of pedestrians and cyclists may change the nature of the traffic flow significantly, or the capacity of a road may vary substantially depending on the volumes of carts and other slow-moving vehicles. Particular capacity

problems are likely to arise at intersections. It may be difficult for drivers to find suitable gaps in the slow stream of pedestrians and cyclists, and consequently the formal rules may frequently be violated. Thus, traffic engineering models or measures from industrialized countries cannot be applied directly. There is a need for careful analysis of the local conditions, including the variety of vehicles and transport modes that are present. In such situations, the road network should be designed to reflect the actual traffic demand, and proper provision should be made for both motorized and non-motorized road users. This may require a fundamental re-evaluation of the underlying assumptions and recalculation of the parameters used in conventional traffic and design models.

3.3 Traffic flows and growth

3.3.1 Traffic

Traffic characteristics

In order to assess benefits in an economic appraisal, it is necessary to separate traffic into the following three categories, since each is treated differently in the cost-benefit analysis (see Chapter 7):

- *Normal traffic*
Traffic which would pass along the existing road if no investment took place, including normal growth.
- *Diverted traffic*
Traffic that changes from another route (or mode) to the road, but still travels between the same origin and destination (this is sometimes termed ‘reassigned traffic’ in transport modelling).
- *Generated traffic*
Additional traffic which occurs in response to the provision or improvement of a road (this includes ‘redistributed traffic’ as sometimes defined in transport models).

Daily flows

The estimate normally used for planning and design is the ‘annual average daily traffic’ (AADT), classified by vehicle category. This is defined as the total annual traffic in both directions divided by 365.

Traffic growth

Estimates must also be made about how this traffic will grow in the future, as many decisions are very sensitive to traffic forecasts. Different methods of forecasting tend to be used depending on the type of traffic being considered. These can be summarized by the following:

- Normal traffic – based on extrapolation of historical time-series data for traffic growth, fuel sales, GNI, or other relevant parameters, but taking into account any specific local circumstances.
- Diverted traffic – estimates normally based on origin-and-destination surveys.
- Generated traffic – normally based on ‘demand relationships’ which indicate the likely increase in traffic level for different levels of cost savings (see Chapter 7); generated traffic is particularly difficult to forecast.

Some methods of forecasting transport demand are discussed in Chapter 6.

3.3.2 Axle loading

The deterioration of pavements caused by traffic results from both the magnitude of the individual axle loads and the number of times that these loads are applied. Factors, such as tyre pressure and wheel configuration, can also be important. For pavement design and maintenance purposes, it is therefore necessary to consider not only the total number of vehicles that will use the road, but also the axle loads of these vehicles. To do this, the axle load distribution of a typical sample of vehicles using a road must be measured. The axle loads can be converted using standard factors to determine the damaging power of different types of vehicle. This damaging power is normally expressed as the number of ‘equivalent standard axles’ (ESA), each of 80 kN, that would do the same damage to the pavement as the vehicle in question. This damaging power is termed the vehicle’s ‘equivalence factor’ (EF) and the design lives of pavements are expressed in terms of ESAs that they are designed to carry.

Impact of axle loading

The relationship between the vehicle’s equivalence factor and its axle loading is normally expressed in terms of the axle mass in kilograms (kg), rather than in terms of the force that they apply to the pavement in kilonewtons (kN). The relationship is normally considered to be of the form:

Damaging power

$$EF = \sum_{i=1}^j \left[\frac{\text{axle}_i}{8160} \right]^n$$

where axle_i is the mass of axle i (kg); j the number of axles on the vehicle in question; n a power factor that varies depending on the pavement construction type, subgrade and assumptions about ‘failure’ criteria, but with a value typically of around 4.0; and the standard axle load is taken as 8,160 kg. This relationship is illustrated in Figure 3.2, where a power factor of 4.0 has been used.

In the past, it has been customary to assume that the axle load distribution of heavy vehicles will remain unchanged for the design life of a new or strengthened pavement. However, more recently it has become clear that there is a tendency for steady growth over time of vehicle axle loads, and forecasts of this are needed to avoid underestimating future pavement damage. There are also examples where the introduction of a fleet of new and different vehicles can radically alter the axle load distribution on a particular route in a short time.

Forecasting

3.4 Capacity and speed–flow

3.4.1 Capacity

As traffic flows increase, average speeds of all vehicles converge towards that of the slowest vehicle in the traffic stream as overtaking opportunities become more and more restricted. The ‘capacity’ of a road is defined as the maximum flow of traffic possible, in unit time, under prevailing traffic and road conditions. As flow approaches capacity, average speeds may fall even lower than the speeds of slow vehicles as a result of small disturbances in the traffic. Flow becomes unstable. ‘Saturation flow’ is defined as the maximum rate of flow possible in the absence of

Saturation flow

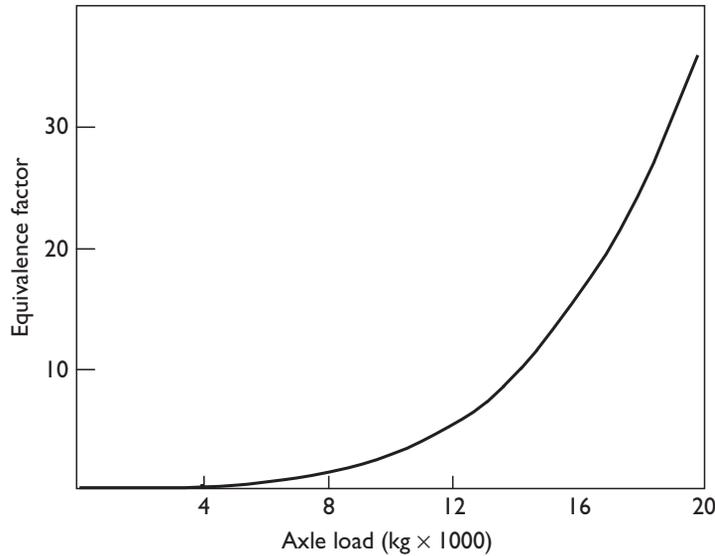


Figure 3.2 Relationship between equivalence factor and axle loading.

controlling factors such as traffic signals. It represents the potential capacity of a road section or junction under ideal conditions.

Passenger car equivalence

Calculations of capacity need to recognize that different types of vehicle take up different amounts of relative space on the road. Thus, traditionally, different vehicle categories have been converted into ‘passenger car units’ (PCUs) by means of a ‘passenger car equivalent’ (E). The value of E is defined as the number of passenger cars that are displaced by a single vehicle of a particular type under the prevailing traffic and road condition. The value does not depend solely on vehicle type. It depends on two main factors: the road condition, such as the type of road and the gradient (a heavy lorry counts more uphill), and the traffic composition, such as the proportion of heavy vehicles or bicycles. Much theoretical and experimental work has been carried out to determine E for different vehicle categories under various road conditions. The value of E is based on measurement of headways between vehicles under saturated flow conditions. The headway of a passenger car following another passenger car is used as the basic value to which other headways are compared.

Examples

A study in Beijing (Zhao 1987) estimated E in mixed traffic with a high proportion of bicycles. A linear regression equation was used to determine E for different vehicle categories. In Beijing, the estimated value of E for bicycles in mixed traffic was found to be 0.18. In a pilot study in Copenhagen, the corresponding factor was found to be between 0.16 and 0.29, depending on the composition of traffic; the values of E for buses and trucks were found to be between 1.54 and 2.21. By comparison, the City Engineer of Copenhagen uses the factors 0.2 for bicycles and 2.0 for buses and trucks. The results indicate that the figure 0.2 may be an adequate estimate for bicycles in most cases. In an Indonesian study (Cuthberth *et al.* 1983), values of E were determined using a linear regression equation with several dependent

variables, such as the width of the road and the speed of traffic. The results for Jakarta are shown in Table 3.5. These results suggest that this method of calculation, using a combination of headway observations and regression analysis, is applicable to third world countries. Other methods of calculating the saturation flow and E are described by Turner (1993).

More recently, the concept of ‘passenger car space equivalents’ (PCSE) has also been used. This accounts only for the relative space occupied by a vehicle on the road. PCSEs are used in situations where traffic models predict explicitly the speed differences of the various vehicles in the traffic stream – such as in *HDM-4* (see Chapter 21). PCSE factors vary by road type, and narrow roads have higher PCSE values than wide roads, as shown in Table 3.6. The concept of PCSEs is described by Hoban *et al.* (1994).

The more complex situations found in urban areas may require the application of a traffic model. In Santiago, Chile, the TRANSYT computer model was used to design fixed-time traffic signal plans. Before the model could be of any use, a comprehensive study was undertaken to determine the basic parameters necessary for calibration (Willumsen and Coeymans 1989). The main parameters analysed were E , the saturation flows, and platoon dispersion parameters, which describe the

PCSE

Traffic models

Table 3.5 Passenger car equivalent (E) by road and carriageway type for Jakarta

Vehicle type	E by road and carriageway type				
	Expressway		Suburban		Urban
	Dual	Dual	Single	Dual	Single
Car, taxi	1.0	1.0	1.0	1.0	1.0
Truck	1.5	1.5	1.7	1.6	2.3
Small truck	1.0	1.0	1.0	1.0	1.0
Large bus	1.8	2.0	2.6	2.4	3.3
Minibus	1.3	1.4	1.8	1.7	2.6
Opelet	1.0	1.1	1.2	1.2	1.7
Three-wheeled vehicles	—	—	0.8	0.8	0.9
Motorcycle	0.7	0.6	0.5	0.5	0.4
‘Becak’	—	—	0.6	0.6	0.6
Bicycle	—	—	0.5	0.4	0.3

Based on data from: Cuthberth (1983).

Table 3.6 Typical PCSE values for different road types

Road type	Width (m)	Capacity (PCSE/h)	Speed at capacity (km/h)
Single-lane	< 4.0	600	10
Intermediate	4.0–5.5	1,800	20
Two-lane	5.5–9.0	2,800	25
Wide two-lane	9.0–12.0	3,200	30
Four-lane	> 12.0	8,000	40

Adapted from: Odoki and Kerali (1999).

way a queue of vehicles released from a stop line is dispersed as a result of differences in acceleration performance and speed.

3.4.2 Variation in flow

30th-hour
volume

It is not economically sensible to design a facility to be congestion-free every hour throughout the year. By defining the distribution of hourly flows over the 8,760 (365×24) hours of the year, the AADT can be converted to hourly flows. Congestion analysis can then be undertaken for a number of hourly flow levels. It is normally considered sound practice to design roads to carry the '30th-hour volume' or 30th HV. This is the 30th-heaviest hour, which is the hourly volume exceeded by only 29 hours per year.

Design hourly
volume

'Design hourly volume' (DHV) is then expressed as $DHV = AADT \times K$, where K is estimated from the ratio of the 30th HV to the AADT from a similar site. The 30th HV, expressed as a fraction of AADT, can vary as indicated in Table 3.7. The higher ratios in Table 3.7 refer to roads with relatively high concentration of traffic during rush hours, or large seasonal variations.

3.4.3 Service volume

Highway
capacity
manual

There is little information from developing and emerging countries concerning traffic capacity in relation to road, traffic and control conditions. Traffic capacity assessments are, therefore, normally based on the *Highway capacity manual* (Transportation Research Board 2000), although this has been developed on the basis of experience gained in the United States. For two-lane asphalt surfaced roads, the *Highway capacity manual* introduces three parameters to describe the service quality: average travel speed, per cent time spent 'following', and capacity utilization. The manual identifies six so-called 'service levels', designated A to F.

Levels of
service

At level A, the highest quality of traffic service occurs and motorists are able to drive at their desired speed of 95 km/h or higher. For ideal conditions (unconstrained geometric, traffic and environmental conditions), a volume of 420 passenger cars per hour, total for both directions, may be achieved. As the level-of-service decreases, so will the average travel speed. The drivers will experience more delays, and platooning becomes intense. At level-of-service E, passing is virtually impossible. At the same time, the vehicle density will increase as vehicles crowd closer and closer together. The highest volume attainable under level-of-service E defines the capacity of the road. Under ideal conditions (no restrictive geometric, traffic or environmental conditions), capacity may reach 2,800 passenger cars per hour, total for both

Table 3.7 Variations in 30th HV

Traffic condition	$K = 30th\ HV/AADT$
Heavily trafficked road under congested rural conditions	0.08–0.10
Normal rural conditions	0.12–0.18
Road catering for recreational or other traffic of seasonal nature	0.20–0.30

Based on data from: SATCC Technical Unit (1995).

directions. However, operating conditions at capacity are unstable and difficult to predict. Level F represents heavily congested flow with traffic demand exceeding capacity. The flow is unstable with unpredictable characteristics. The six levels of service are presented in the speed–flow diagram, shown in Figure 3.3.

Guidelines for selection of levels of service for the design of different classes of rural roads are given in Table 3.8.

The capacity of two-lane roads varies, depending on the terrain and the extent of passing restrictions. Volume–capacity (V/C) ratios are given in Table 3.9 for different levels of service, relative to the ‘ideal capacity’ of 2,800 passenger cars per hour (flat terrain and zero no-passing zones).

The service volume, SV , is defined as the maximum volume of traffic that the road is able to carry under the prevailing conditions, without falling below the preselected level of service. The traffic data required for determining the service volume include the two-way hourly volume, the directional distribution of traffic flow and the

Selecting levels of service

Volume–capacity ratios

Determining the service volume

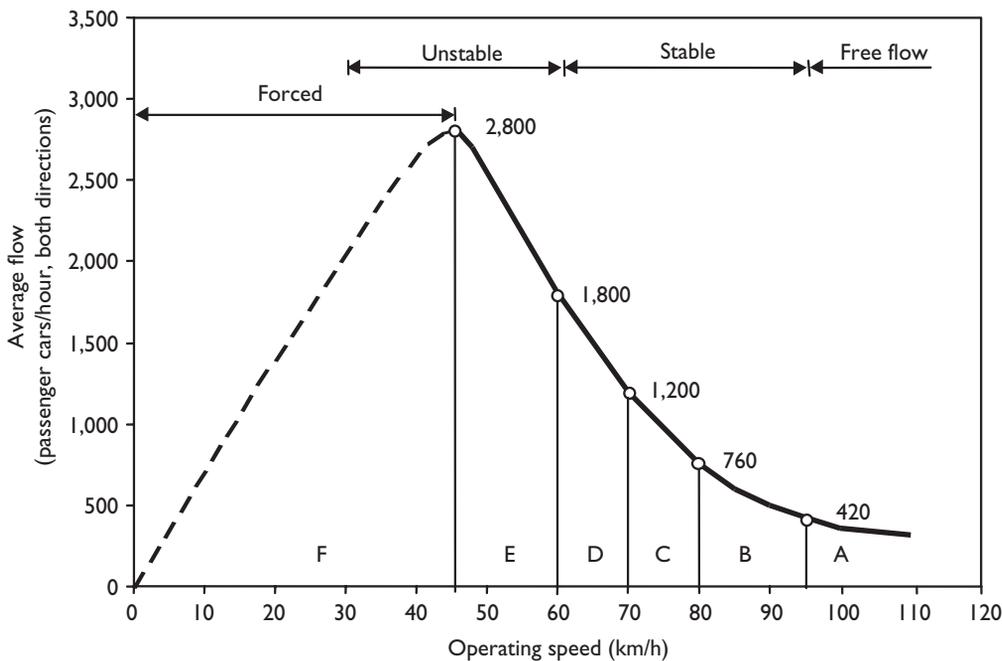


Figure 3.3 Speed–flow diagram.

Adapted from: SATCC Technical Unit (1995).

Table 3.8 Level of service for rural road design

Road class	Level of service		
	Flat terrain	Rolling terrain	Mountainous terrain
Arterials	B	B	C
Collectors	C	C	D
Local roads	D	D	D

Source: SATCC Technical Unit (1995).

Table 3.9 Volume–capacity ratio

Level of service	Per cent time spent following	Per cent no-passing zones ^a	Volume–capacity ratio (VIC) ^b (total in both directions)		
			Flat terrain	Rolling terrain	Mountainous terrain
A	<30	0	0.15	0.15	0.14
		20	0.12	0.10	0.09
		40	0.09	0.07	0.07
B	30–45	0	0.27	0.26	0.25
		20	0.24	0.23	0.20
		40	0.21	0.19	0.16
		60	0.19	0.17	0.13
C	45–60	0	0.43	0.42	0.39
		20	0.39	0.39	0.33
		40	0.36	0.35	0.28
		60	0.34	0.32	0.23
		80	0.33	0.30	0.20
D	60–75	0	0.64	0.62	0.58
		20	0.62	0.57	0.50
		40	0.60	0.52	0.45
		60	0.59	0.48	0.40
		80	0.58	0.46	0.37

Source: SATCC Technical Unit (1995).

Notes

a The percentage of road length where the sight distance is less than 450 m may be used as per cent no-passing zones.

b Ratio of hourly volume to an ideal capacity of 2,800 passenger cars per hour in both directions.

proportion of trucks and buses in the traffic stream. For paved two-lane rural roads, SV can be determined from the general relationship:

$$SV_i = 2800 (V/C)_i f_W f_T f_D$$

where SV_i is the total service volume for level of service 'i' in vehicles per hour; $(V/C)_i$ the volume–capacity ratio for level of service 'i'; f_W the lane width factor; f_T the truck factor; and f_D the directional factor. The adjustment factors, f_W , for the combined effect of narrow lanes (less than 3.5 m) and restricted shoulder width, are given in Table 3.10. The truck factor, f_T , is dependent on the level of service, the type of vehicle and the type of terrain. It can be determined from the relationship

$$f_T = \frac{1}{1 + P_T(E_T - 1) + P_B(E_B - 1)}$$

where P_T is the proportion of trucks in the traffic stream, expressed as a decimal; P_B the proportion of buses in the traffic stream, expressed as a decimal; E_T the passenger car equivalent for trucks, and E_B the passenger car equivalent for buses, obtained from Table 3.11.

Table 3.10 Lane width factor, f_w

Usable shoulder width (m)	Lane width (m)			
	3.5	3.35	3.00	2.75
> 1.8	1.00	0.93	0.84	0.70
1.2	0.92	0.85	0.77	0.65
0.6	0.81	0.75	0.68	0.57
0	0.70	0.65	0.58	0.49

Source: SATCC Technical Unit (1995).

Table 3.11 Passenger car equivalents

Vehicle type	Level of service	Type of terrain		
		Flat	Rolling	Mountainous
Trucks (E_T)	A	2.0	4.0	7.0
	B and C	2.2	5.0	10.0
	D and E	2.0	5.0	12.0
Buses (E_B)	A	1.8	3.0	5.7
	B and C	2.0	3.4	6.0
	D and E	1.6	2.9	6.5

Source: SATCC Technical Unit (1995).

Table 3.12 Directional factor, f_D

Directional distribution	50/50	60/40	70/30	80/20
Directional factor, f_D	1.00	0.94	0.89	0.83

Source: SATCC Technical Unit (1995).

The V/C values given in Table 3.9 are for 50/50 directional distribution of traffic. As directional split moves away from the ideal 50/50 condition, the capacity is reduced by the factor f_D as shown in Table 3.12.

The maximum AADT for two-lane rural roads, applicable to different levels of service and different types of terrain, is given in Table 3.13. In this table, the AADT is based on an assumed traffic mix of 25 per cent trucks, a 60/40 directional split and 20, 40 and 60 per cent no-passing zones for flat, rolling and mountainous terrain respectively. A ratio of 30th HV to AADT of 0.15 has been assumed for normal rural conditions, based on Table 3.7.

3.4.4 Gravel roads

For gravel roads, the presence or absence of dust and the roughness of the running surface have a significant influence on the service volume. Although gravel roads

Directional distribution

Maximum flows

Table 3.13 Maximum AADT for two-lane rural roads

Level of service	Maximum AADT		
	Flat terrain	Rolling terrain	Mountainous terrain
A	1,600	700	300
B	3,200	1,650	700
C	5,200	3,000	1,250
D	8,700	4,500	1,900

Source: SATCC Technical Unit (1995).

are rarely improved for capacity reasons, the nominal service volume of a two-lane gravel road can be assumed to be around 10 per cent of that of an asphalt surfaced road with the same alignment and cross-sectional characteristics. Improvement to asphalt surfaced standard for economic reasons (maintenance and road user cost savings) normally takes place at much lower volumes than is warranted by the capacity.

3.5 Increasing capacity

3.5.1 Geometric improvement

Traffic separation

One of the most efficient ways of improving the capacity of a road is to separate motor vehicles from slow moving traffic such as pedestrians, bicycles and carts. The capacity of an intersection approach in Beijing is compared with mixed traffic flow and a separate bicycle track in Table 3.14. This shows that the total capacity of the approach increases when bicycle traffic is separated from motor traffic. These results do not only apply in cities. In rural areas, a combined pedestrian and bicycle track beside the road will improve both the capacity and the safety of the road.

Intersections

The above findings relate to road sections and approaches to intersections. Problems can also arise within intersections because of the conflicts between different categories of traffic. Often, large proportions of pedestrians and cyclists obstruct the intended flow of motorized traffic – and vice versa. The most efficient countermeasure is to design the intersections with a large degree of channelization. This separates and directs traffic, and provides waiting areas for pedestrians. Separation is obtained by the use of turning lanes and waiting areas, raised traffic islands, pedestrian refuges and carriageway markings. Even a simple pedestrian crossing will benefit from a central island, allowing pedestrians to cross the road in two stages, and at the same time slowing down traffic and preventing overtaking adjacent to the crossing.

Self-enforcing measures

The aim of many traffic engineering measures, such as channelization, is to facilitate the drivers' task by providing necessary information, and by reducing conflicts and the choices to be made. In many industrialized countries, there is an increasing recognition of the importance of designing the road, its equipment and surroundings to be 'self-enforcing', with the aim of simplifying driving. This approach is even

Table 3.14 Comparison of the capacity of an approach with mixed respectively separated traffic

Traffic type	Motor traffic (PCU/h)	Bicycles (per hour)	Total flow (PCU/h)
Mixed traffic	754	4,184	1,507
Separated traffic	965	4,320	1,743

Based on data from: Zhao (1987).

more relevant to developing and emerging countries, where the traffic environment and attitudes of drivers are still developing.

3.5.2 Public transport

A large proportion of trips are made by public transport. Consequently, improving the conditions under which public transport operates can have a significant impact on the transport system more widely (World Bank 1986). The challenge is to introduce traffic engineering measures that are self-enforcing and to facilitate improved behaviour of bus drivers and passengers.

Segregation of traffic streams, with separate bus lanes or busways, can improve traffic flow conditions considerably. Priority treatment of buses in signal-controlled intersections can improve travel speed. Studies of bus lanes in Bangkok (Marler 1982) concluded that not only bus travel speed increased but, more surprising, car travel speed was also improved. The banning of kerbside parking along bus lanes contributed to this. Traffic operations can also be improved significantly by reserving separate streets or parts of the street as a comprehensive busway network (Gardner *et al.* 1991).

An important issue in relation to public transport is the location and design of bus stops. The stops must be easily accessible for both passengers and buses. Passengers should be able to approach bus stops without crossing traffic lanes, as is often the case with streetcars or bus lanes located in the centre of the road. Bus stops should have waiting areas large enough to accommodate passengers without any danger of being hit by passing vehicles (Marler 1982). Larger bus stations should have an easily understood information system to direct the passengers to the appropriate bus stop.

3.6 Traffic information and data

Information on traffic is needed to support a wide range of decisions relating to roads. Some traffic data may be readily available, but this may be limited to the number of registered motor vehicles in different classes, published as national statistics. In most cases, traffic surveys or other studies will need to be undertaken. A large number of different survey types can be used, depending on data needs, and those applicable to urban roads are listed in Table 3.15. Clearly, a large number of survey types is possible, and reference should be made to *Overseas Road Note 11* (TRL 1993)

Aim

Priority

Bus stops

Traffic
information

Table 3.15 Types of urban traffic surveys

Type	Information	Method	Output
Road inventory	Road network characteristics	Observation	<ul style="list-style-type: none"> • Geometry • Land-use • Road-furniture provision
Parking inventory	Parking supply	Observation	<ul style="list-style-type: none"> • Available parking space • Types of parking
Origin–destination	Demand forecasting	Registration number method	<ul style="list-style-type: none"> • Route choice • Through-traffic • Travel times
Traffic volumes	Demand	Manual counts, automatic counts	<ul style="list-style-type: none"> • Vehicle flows on links • Junction movements • Passenger flows • Traffic variability • Peak-hour factors • AADT
Spot speeds	Vehicle performance on links	Short-base method, radar observation	<ul style="list-style-type: none"> • Vehicle speeds on links • Speed–flow measurements
Network speeds and delays	Route network performance	Floating car method	<ul style="list-style-type: none"> • Network speeds • Link speeds • Network delay • Congestion points
Junction delay	Junction performance	Stopped vehicle count, elevated observer method	<ul style="list-style-type: none"> • Total delays • Average delays on arms • Distribution of delay times by turning movement • Delay causes
Saturation flows	Junction capacity	Flow profile method, saturated period count	<ul style="list-style-type: none"> • Saturation flow • Junction capacity

Source: TRL (1993).

for more details of these. Data issues relating to some of these survey types are now described.

Planning

Before a survey is carried out, comprehensive planning needs to be undertaken to ensure that the required results are obtained. The need for information must be considered carefully and limited with respect to the cost involved in collecting the data. The World Bank has devised a system of ‘information quality levels’ (IQL) for assisting with determining data requirements (Paterson 1991). Four levels of information are used, and examples for traffic volume and axle load data are given in Table 3.16. Robinson *et al.* (1998) have proposed that different IQLs should be used in connection with collecting and using data for different types of road management operations (see Chapter 19):

- *IQL-IV* – strategic planning (long-term, network-wide information)
- *IQL-III* – programming (medium-term, budget preparation)
- *IQL-II* – preparation (detailed design and contract preparation)
- *IQL-I* – detailed operations management and special studies.

Table 3.16 Traffic volume and axle loading data for different IQL values

Data group	Information quality level			
	I	II	III	IV
<i>Volume</i>				
Total volume	AADT, seasonal ADT, hourly and short-term flows	AADT and seasonal AADT	AADT and seasonal factor	AADT range
Directional characteristics	By direction and lane	By direction, average heavy vehicles per lane	None	None
Composition	By vehicle class	By vehicle class	By 2–3 categories (e.g. heavy, bus, light)	Proportion of heavy vehicles
<i>Loading</i>				
Axle loading	Axle load spectrum	Average ESA per vehicle class, maximum axle load	Link/region average ESA per vehicle class	Regional average ESA per heavy vehicle
Gross vehicle mass	Spectrum by vehicle class	Average and maximum by class	None	None
Tyre pressure	Average and maximum by vehicle class	None	None	None

Source: Paterson (1991).

Whatever type of data or level of information being collected, it is important to plan for the processing of the data well ahead of the survey. The need for information, the type of survey, the survey forms and the data processing, analysis and presentation of data must form a coherent whole in order not to waste valuable resources.

An example of traffic data collected is given in Table 3.17. This shows data on the modal share of motorized trips for different types of vehicle in a number of major cities. Note that these data provide no information about pedestrian and bicycle trips, so could only be used for studies where this information was not relevant.

Traffic studies on rural roads often need to measure current flows and then to predict future growth. Thus, the need is to estimate the ‘baseline’ traffic flows or, in other words, to determine the traffic volume actually travelling on the road at present. Estimates of AADT are normally obtained by recording actual traffic flows over a specific shorter period than a year, and results are scaled to give an estimate of AADT. Both manual and automatic methods of counting can be used for this, and each is appropriate in different situations. Both methods are prone to inaccuracy: automatic methods because of difficulties of setting vehicle sensors to record vehicles or axles correctly, and manual methods because of human error.

Traffic counts carried out over a short period, as a basis for estimating the AADT, can produce estimates that are subject to large errors, because traffic flows can have

Example

Traffic on rural roads

Uncertainty

Table 3.17 Examples of modal share for motorized trips only (per cent)

City	Cars	Taxis	Buses	Paratransit	Rail or metro	Other
Abidjan	33	12	50	—	—	5
Amman	44	11	19	26	0	0
Ankara	23	10	53	9	2	2
Bangkok	25	10	55	10	—	—
Bogotá	14	1	80	0	0	5
Bombay	8	10	34	13	34	—
Buenos Aires	—	—	45	27	—	28
Cairo	15	15	70	—	—	—
Calcutta	—	2	67	14	10	4
Hong Kong	8	13	60	—	19	—
Jakarta	27	—	51	—	1	21
Karachi	3	7	52	18	6	13
Kuala Lumpur	37	—	33	17	0	13
Lima	—	—	45	27	—	28
Manila	16	2	16	59	—	8
Mexico City	19	—	51	13	15	2
Nairobi	45	—	31	15	0	9
Sao Paulo	32	3	54	—	10	1
Tunis	24	4	61	—	10	—
London	61	1	23	0	12	2
New York	12	2	14	0	72	0
Paris	56	—	8	0	21	15
Tokyo	32	—	6	0	61	0

Adapted from: World Bank (1986).

large hourly, daily, weekly, monthly and seasonal variations. The daily variability in traffic flow depends on the volume of traffic, increasing as traffic levels fall, and with high variability on roads carrying less than 1,000 vehicles per day, as shown in Figure 3.4. Traffic flows vary more from day-to-day than week-to-week over the year, so there are large errors associated with estimating annual traffic flows (and subsequently AADT) from traffic counts of a few days duration. For the same reason, there is a rapid fall in the likely size of error as the duration of counting period increases up to one week, but there is a marked decrease in the reduction of error for counts of longer duration, as shown in Figure 3.5. Traffic flows also vary from month-to-month, so that a weekly traffic count repeated at intervals during the year provides a better base for estimating the annual volume of traffic than a continuous traffic count of the same total aggregate period. Traffic also varies considerably through the day, but this is unlikely to affect the estimate of AADT providing sufficient and appropriate hours are covered by the daily counts. The key issue is that, recognizing that estimates of traffic flow are subject to large variability, the amount of effort put in to reducing the variability should depend on the use to which the data will be put.

Difficulty of forecasting growth

Forecasting growth reliably is notoriously difficult, especially when considering the variability likely to be present in estimates of current flows, as discussed above. Even in industrialized countries with stable economic conditions, large errors can

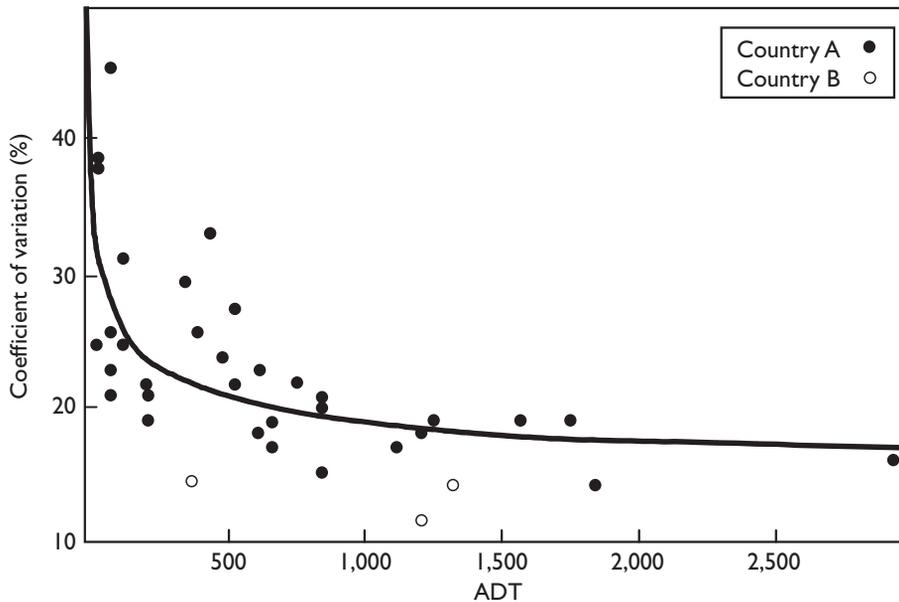


Figure 3.4 Relationship between daily variability and traffic flow.

Adapted from: Howe (1972).

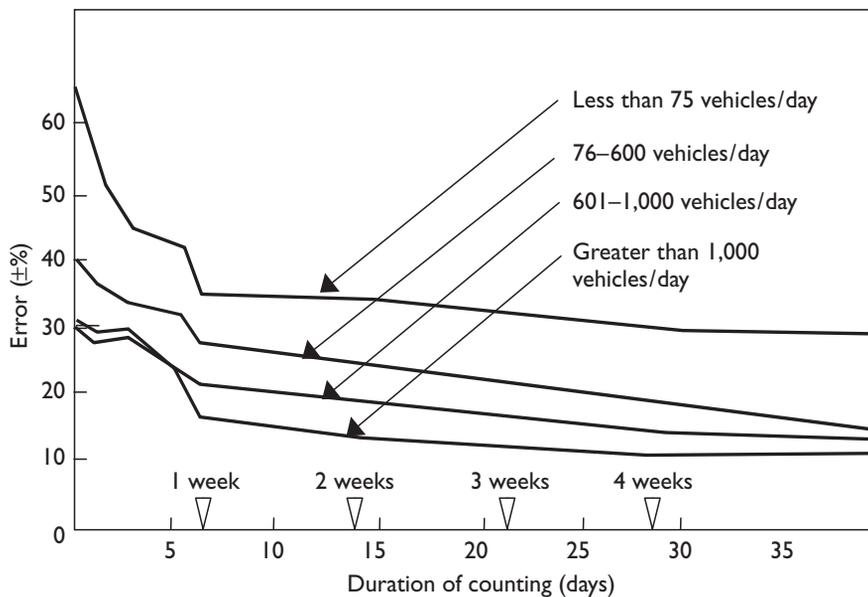


Figure 3.5 Errors in AADT estimates from random counts of varying duration.

Adapted from: Howe (1972).

occur but, in countries with developing or transitional economies, the problem becomes more intractable.

The only effective way to determine the damaging effect of traffic on different roads is to measure the complete spectrum of axle loads, and to calculate

Axle load surveys

the consequential damage in terms of standard axles. Axle load surveys can be carried out using a variety of types of equipment, ranging from small portable weighbridges, through 'weigh-in-motion' systems, to fixed static weighing platforms. Changes in future axle loading cannot normally be forecast and, hence, extrapolation from existing axle load surveys cannot provide for this sort of eventuality. Regular traffic surveys should highlight such situations, and new axle load surveys should then be carried out, as appropriate.

References

- Cuthberth, R. *et al.* (1983) 'The Jakarta traffic management study, 2: assessment of PCU values', *Traffic Engineering and Control*, **24**(8).
- Gardner, G. *et al.* (1991) 'The performance of busway transit in developing cities', *TRRL Research Report 329*, Crowthorne: Transport and Road Research Laboratory.
- Habitat (1984) *Transportation Strategies for Human Settlements in Developing Countries*, Nairobi: United Nations Commission for Human Settlements.
- Hoban, C., Reilly, W. and Archondo-Callao, R. (1994) 'Economic analysis of road projects with congested traffic', in *Transportation, Water & Urban Development Department, Methods of Economic Evaluation of Highway Investments and Maintenance*, Washington, DC: The World Bank.
- Howe, J.D.G.F. (1972) 'A review of rural traffic-counting methods in developing countries', *RRL Report LR427*, Crowthorne: Road Research Laboratory.
- Marler, N.W. (1982) 'The performance of high-flow bus lanes in Bangkok', *TRRL Supplementary Report 723*, Crowthorne: Transport and Road Research Laboratory.
- Odoki, J.B. and Kerali, H.R. (1999) 'HDM-4 analytical framework and model descriptions', *The Highway Development and Management Series, Volume 4*, Paris: PIARC World Road Association.
- Paterson, W.D. (1991) 'Choosing an appropriate information system for road management, in *PIARC 19th World Congress, Marrakesh, September 1991*, Paris: Permanent International Association of Road Congresses.
- Robinson, R., Danielson, U. and Snaith, M. (1998) *Road Maintenance Management: Concepts and Systems*, Basingstoke: Macmillan.
- SATCC Technical Unit (1995) *Recommendations on Road Design Standards, Volume I: Geometric Design of Rural Roads*, Maputo: Southern Africa Transport and Communications Commission.
- Transportation Research Board (2000) *Highway Capacity Manual 2000*, Washington, DC: National Academy of Sciences.
- TRL (1993) 'Urban road traffic surveys', *Overseas Road Note 11*, Crowthorne: Transport Research Laboratory.
- Turner, J. (1993) 'Simplified saturation flow data collection methods', *TRL Project Report PR/OSC/003/93*, Crowthorne: Transport Research Laboratory.
- Willumsen, L.G. and Coeymans, J.E. (1989) 'The value of fixed-time signal co-ordination in developing countries, II: improved bus modelling and results', *Traffic Engineering and Control*, **30**(3).
- World Bank (1986) 'Urban transport', *World Bank Policy Study*, Washington, DC.
- Zhao, Song (1987) 'Capacity of signalized intersections: a study carried out in Beijing, China', *Institute of Roads, Transport and Town Planning, Report no 53*, Lyngby: Technical University of Denmark.

Traffic safety

Goff Jacobs and Chris Baguley

4.1 Background

The problem of deaths and injury as a result of road accidents is now acknowledged to be a global phenomenon, with authorities in virtually all countries concerned about the growth in the number of people killed and seriously injured on their roads. In recent years, two major studies of causes of death worldwide have been published: *The Global Burden of Disease* (Murray and Lopez 1996) and the *World Health Report – Making a Difference* (WHO 1999). These show that, in 1990, road accidents lay in ninth place out of a total of over hundred separately identified causes of death and disability. However, the reports also forecast that, by the year 2020, road accidents will move up to sixth place and, in terms of ‘years of life lost’ (YLL) and ‘disability-adjusted life years’ (DALYs – which express years of life lost to premature death, and years lived with disability, adjusted for the severity of the disability), will be in second and third place respectively.

Global
problem

It is now estimated that between 750,000 and 800,000 people are killed each year in road accidents globally – if the under-reporting of road accidents, particularly in the developing world, is taken into account (Jacobs *et al.* 2000). About 70 per cent occur in those countries classified as low- or middle-income, with a further 18 per cent from the Middle Eastern, Central and Eastern European countries. Thus, less than 14 per cent of global road deaths occur in the so-called developed regions of the world, even though they contain almost two-thirds of the world’s motor vehicles. A conservative estimate of global injuries suggested that, in 1999, up to 34 million people were injured worldwide in road accidents, again with about 75 per cent occurring in the developing and emerging countries.

Numbers
killed and
injured

4.2 Nature of the problem

4.2.1 Rates and trends

The usual basis of comparison of the seriousness of the road accident problem in different countries is the number of deaths from road accidents per annum per 10,000 vehicles licensed. However, this is far from ideal as an indicator, since it takes no account of road user exposure to risk. For example, the injury accidents per million vehicle-kilometres travelled per annum would be more appropriate, but the reporting of non-fatal accidents in most developing and emerging countries is poor, and few carry out surveys on annual travel by different classes of vehicle.

Basis of
comparison

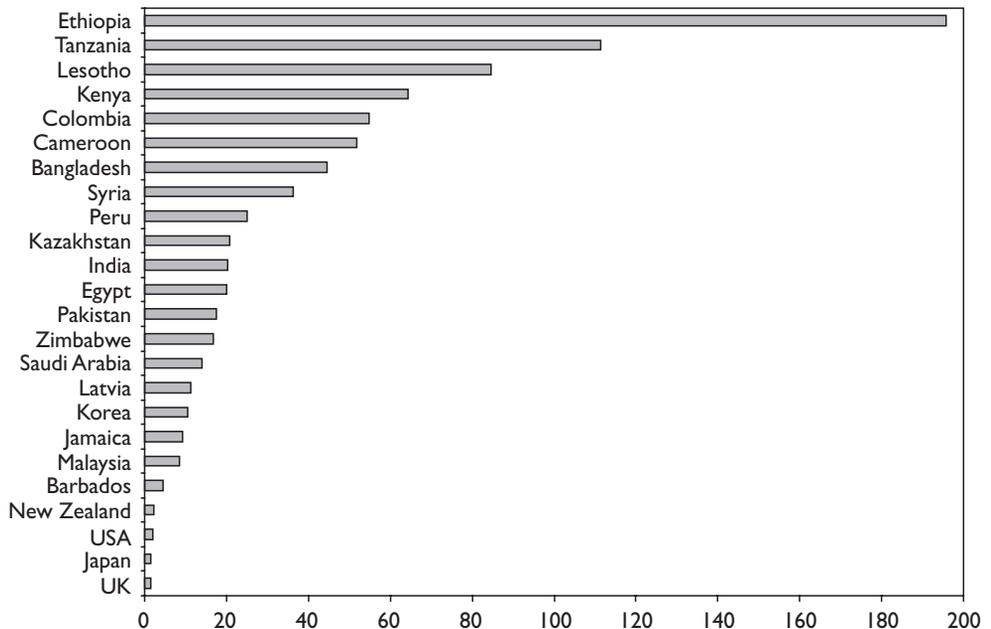


Figure 4.1 Fatalities/10,000 licensed motor vehicles in selected countries (1996 or latest year available).

Comparison of rates

Figure 4.1 shows that countries of Western Europe and North America are characterized by a death rate of often less than 2 per 10,000 licensed vehicles. However, some developing countries have death rates in excess of 150. In most developing and emerging countries, road accident deaths are under-reported and licensed vehicles are overestimated because scrapped vehicles tend not to be removed from the vehicle register. Both of these factors acting together suggests that actual fatality rates might well be greater than those shown in Figure 4.1.

Examples of under-reporting

A study in Colombo, Sri Lanka, compared 'official' road accident statistics from police records with those held by hospitals. It was found that fewer than 25 per cent of the hospital records of fatal and serious accidents were contained in the police data. A further study in selected cities found wide variation, with estimates of police recording between 3 per cent (in Hanoi) and 57 per cent (in Harare) of the actual number of casualties. Matching of accidents involving children and women was particularly low. Similarly, in the Philippines, only 1 out of 5 medically reported road deaths are included in police statistics. In Indonesia, insurance companies report almost 40 per cent more deaths than the police. The Department of Health in Taiwan, in 1995, reported 130 per cent more deaths than the police. In Karachi, ambulance statistics showed that only about half of road accident deaths were reported by the police. Under-reporting also appears to be high in China, which already has the world's highest number of road deaths. In 1994, it was estimated that the actual number of people killed in road accidents was about 111,000, which is over 40 per cent greater than the 78,000 reported officially by the police. Thus, the overall safety problem in developing countries is likely to be much worse than official statistics suggest.

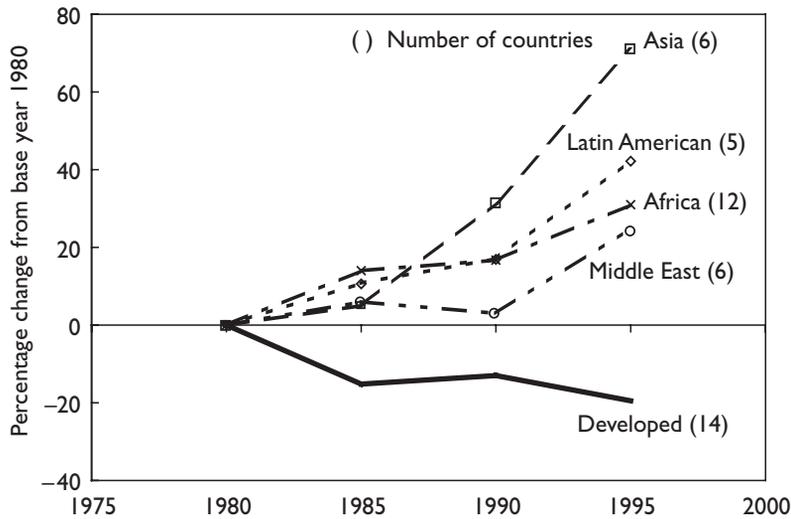


Figure 4.2 Regional trends in road deaths 1980–95.

Figure 4.2 shows the percentage increase or decrease in the actual number of road accident fatalities over the period 1980–95 for five groups of countries. It can be seen that, during this period, the number of road accident deaths in 14 developed countries actually fell on average by 20 per cent. Conversely, in the sample of Asian, Latin American, African and Middle Eastern countries, for which reasonably accurate statistics were available, there were increases ranging from 25 to over 71 per cent.

Change over time

4.2.2 Accident patterns

Some common accident characteristics in developing and emerging countries differ from those found in industrialized countries, as shown in Figure 4.3 and Table 4.1. For example, a relatively high proportion of fatalities are pedestrians and children aged under 16 years, and many fatal accidents involve trucks, buses and other public service vehicles. In many cases, these higher percentages are a consequence of the differences between the traffic and population characteristics. For example, the average percentage of the population aged 5–14 years in a sample of 16 developing countries was 28 per cent compared with 15 per cent for 9 industrialized countries (Downing and Sayer 1982). As pedestrians, children and professional drivers constitute such a large proportion of the accident problem, it is clear that in many developing and emerging countries road safety engineers and planners need to give priority to improving the safety of these particular groups.

4.3 Road accident costs

In addition to the humanitarian consequences of reducing road deaths and injuries, a strong case can be made for reducing accidents solely on economic grounds, as they consume massive financial resources that countries can ill afford to lose. Even within the transport and road sector, hard decisions have to be taken on the resources

Need for cost information

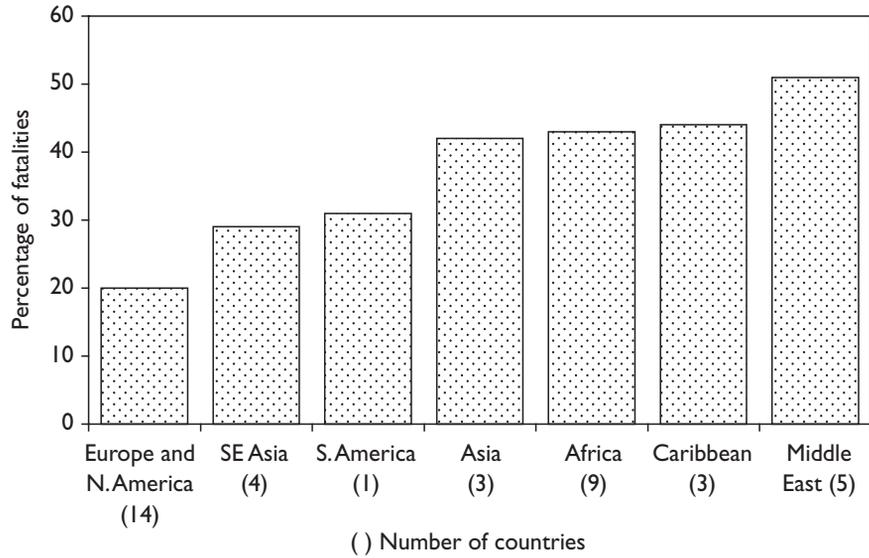


Figure 4.3 Pedestrian fatalities as a percentage of all road accident fatalities.

Table 4.1 Characteristics of fatal accidents

	Percentage of fatalities involving	
	Children under 16 years	Trucks and buses
Botswana	16	25
Egypt	12	37
Ghana	28	50
Pakistan (Karachi)	14	44
Guinea	20	37
Zimbabwe	11	45
United Kingdom	9	21

that a country can devote to road safety. However, in developing and emerging nations, road safety is but one of the many problems demanding its share of funding and other resources. It is therefore essential that a method be devised to determine the cost of road accidents and the value of preventing them.

Level of detail

The first need for costs is at the level of national resource planning to ensure that road safety improvements receive their fair share of investments. Broad estimates are usually sufficient for this purpose, but the level of accuracy needs to be compatible with competing sectors. A second need for costs is to ensure that the most appropriate investments in safety improvements are undertaken in terms of costs and benefits. Failure, by engineers and planners, to associate specific costs with road accidents will result in a lack of consistency in the choice of remedial measures and the assessment of road safety projects. As a consequence, investments in road safety are unlikely to be 'optimal'. In particular, if safety benefits are ignored in transport planning, then there will be underinvestment in road safety.

A number of different methods are available for costing road accidents, and some of these are summarized in Box 4.1. An analysis of 21 studies worldwide – 1 in Latin America, 7 studies in Asia, 4 in Africa, 1 in the Middle East and 8 in industrialized countries (Jacobs *et al.* 2000) showed that all developing countries used the ‘human capital’ approach to costing accidents, while the majority of industrialized countries used the ‘willingness to pay’ approach.

Although there is no real substitute, in individual countries, to carrying out a detailed appraisal of national accident costs, expressing accident costs as a percentage of gross national income (GNI) provides a useful, albeit crude, approach to costing accidents, particularly on a global or regional basis. A study conducted almost a quarter of a century ago estimated road accidents to cost on average 1 per cent of a country’s GNI. This figure has been used by many countries and international aid agencies to estimate the scale of costs incurred by road accidents. As countries have

Costing
methods

GNI and
global costs

Box 4.1 Accident costing methods

Gross output (or human capital) approach

This assesses costs that are due to a loss of current resources and also those due to a loss of future output. In variants of this approach, a significant sum is added to reflect the ‘pain, grief and suffering’ of the victim, and those who care for him or her.

Net output approach

This is similar to the gross output method but, in this case, the discounted value of the victim’s ‘future consumption’ is subtracted from the gross output figure.

Life insurance approach

Costs are related to the sums for which ‘typical’ individuals are willing to insure their own life (or limbs).

Court award approach

Sums awarded by the courts to dependants of those killed or injured provide an indication of the ‘cost’ of the accident.

Implicit public sector approach

Costs are determined by using values that society places on accident prevention in safety legislation, or in public sector decisions taken either in favour or against investment programmes that affect road safety.

Value of risk change or willingness to pay approach

Value is defined in terms of the aggregate amount that people are prepared to pay for a road safety improvement. Conversely, the cost of a reduction in safety is defined in terms of the amount people would require in compensation for the increased risk presented.

Adapted from: *Overseas Road Note 10* (TRL 1995).

Table 4.2 Road accident costs by region (US\$ billion) 1997

Region	Estimated annual accident costs		
	Regional GNI	Per cent GNI	Cost
Africa	370	1.0	3.7
Asia	2,454	1.0	24.5
Latin America/Caribbean	1,890	1.0	18.9
Middle East	495	1.5	7.4
Central and Eastern Europe	659	1.5	9.9
Highly motorized countries	22,665	2.0	453.0
Total			518.0

developed, a higher range of 1–3 per cent has been suggested by the World Bank and others as being more appropriate. Table 4.2 provides an estimate of global and regional costs based on assumptions about relationships with GNI. This table implies that road accident costs may be of the order of US\$65 billion in developing and transitional countries, and US\$453 billion in highly motorized countries, with a total of US\$518 billion worldwide.

Specific country cost analyses

A study carried out of national accident costs in different countries (Jacobs *et al.* 2000) expressed these as a percentage of GNI. Results ranged from 0.3 per cent in Vietnam, and 0.5 per cent in Nepal and Bangladesh, to almost 5 per cent in the United States, Malawi and Kwa Zulu Natal in South Africa. It should be noted that, in this analysis, the costs determined by the different countries have been used directly and not amended in any way. However, relatively little is known about the accuracy of the costing procedures used in each country; for example, whether or not under-reporting of accidents has been taken into account; how damage-only accidents have been assessed; what sums (if any) have been added to reflect pain, grief and suffering; if the human capital approach has been used, etc. Overall, it does appear that in most countries, costs exceed 1 per cent of GNI, which may now be considered to be an underestimate of national accident costs. However, the figures also indicate that costs as a percentage of GNI may be lower in less developed countries.

4.4 Contributory factors

Sources of data

In most countries, police road accident reports give some information about the factors or causes that contribute to accidents. In general, these data should be treated with some caution by those concerned with the introduction of safety measures. The police investigating the accidents are unlikely to have been trained as engineers, and they may therefore underestimate the contribution made by road engineering problems. Their main aim is usually to determine whether or not there has been a traffic violation. The emphasis of the investigation is likely to be placed on detecting human error and apportioning blame. In the United Kingdom, in the early 1970s, 'on-the-spot' investigations were carried out in an area of South East England (Sabey and Staughton 1975). This study demonstrated the importance of road-user factors, which contributed to 95 per cent of the accidents. There was also a strong link

Table 4.3 Causes of road accidents as determined by the police in different countries

	Main cause of accident (per cent)			
	Road-user error	Vehicle defect	Adverse road conditions or environment	Other
Afghanistan	75	17	9	—
Botswana	94	2	1	3
Cyprus	94	1	6	—
Ethiopia	81	5	—	14
India	80	7	1	12
Iran	64	16	20	—
Pakistan	91	4	5	—
Philippines	85	8	7	—
Malaysia	87	2	4	7
Zimbabwe	89	5	1	5
United Kingdom ^a	95	8	28	—

Note

^a Results of on-the-spot studies; note that, in about 30 per cent of accidents, multiple factors were identified (hence totals can add to greater than 100 per cent).

between road-user error and deficiencies in the road environment, which together contributed to over 25 per cent of accidents (see Table 4.3). These results have proved extremely useful in providing an impetus for road safety engineering work undertaken since that time. Similar studies have also been carried out in the United States and Australia, and a new on-the-spot study was begun in the United Kingdom in 2000.

Constraints of expertise or funding have prevented studies of this type being carried out in developing countries, so police reports are the only source of information available. The data from Table 4.3 highlight the importance of road-user error as a factor in all developing countries, but give little indication of any road environment factor, other than in the case of Iran. It seems likely that the importance of road environment factors has been underestimated considerably by the police in their statistics, and this should be borne in mind by engineers and planners. The condition of main roads tends to be poorer in developing than in developed countries (see, e.g. Harral and Faiz 1988) and the pace of introducing engineering improvements to reduce road accidents is considerably slower.

Studies of road-user behaviour (Jacobs *et al.* 1981) at traffic signals and pedestrian crossings in a number of third world cities indicated that road users tended to be less disciplined than, for example, in the United Kingdom. Observations in Pakistan have demonstrated relatively high proportions of drivers crossing continuous 'no-overtaking' lines (15 per cent) and not stopping at stop signs (52 per cent). Although the relationship between these differences in behaviour and accidents has not been clearly established, the results suggest that road safety measures that are not self-enforcing, such as road signs and markings, may be much less effective unless they are integrated with publicity and enforcement campaigns. Poor

Developing and emerging countries

Road-user behaviour and knowledge

road-user behaviour exhibited by drivers in some developing and emerging countries may be due to their lack of knowledge about road safety rules and regulations, or their general attitude towards road safety matters. A study of drivers' knowledge in Jamaica, Pakistan and Thailand indicated that there were only a few topics where a lack of knowledge was widespread. One such example was stopping distances, where 87 per cent of the drivers underestimated the distance required to stop in an emergency when travelling at 50km/h. Professional drivers in Cameroon and Zimbabwe were unable to answer more than half the questions on driving knowledge and skills correctly, and questions on stopping and following distances proved particularly problematic. Other areas of driver behaviour, such as not stopping at pedestrian crossings, traffic signals and stop signs, were found to be due to poor attitudes rather than to poor knowledge. Although attitudes are notoriously difficult to change, there would seem to be some potential for improving them by introducing publicity and enforcement campaigns.

Alcohol

Another area of concern in some, but not all, countries is the problem of alcohol and road users. For example, research from roadside random breath test surveys in Papua New Guinea has indicated that the proportion of weekend drivers sampled after 10:00 p.m. with over 80 mg alcohol/100 millilitres blood in their bodies is more than ten times that found in the United Kingdom (Hills *et al.* 1996). Indian studies have indicated that, on certain roads after sunset, between 50 and 65 per cent of accidents are due to drinking and driving.

Technology transfer

There appear to be wide differences between developed and developing countries in the behaviour, knowledge, attitudes and culture of the road users, in the conditions of the roads and the vehicles, and in the characteristics of traffic. Consequently, the effectiveness of transferring some developed country solutions to developing and emerging countries is uncertain, and their appropriateness needs to be considered by engineers and planners in relation to the problems and conditions prevailing in individual countries.

4.5 Institutions and information systems

4.5.1 Institutional requirements

Multi-sectoral co-ordination

The process of planning and implementing road safety improvements needs to be multi-disciplinary and dynamic. Co-ordination between the various bodies involved in road safety activities, such as the engineers, police, the health sector is essential. There will often be a need to strengthen the institutions responsible for the various aspects of road safety, and to increase their capability for multi-sectoral action. However, there is no single 'best way' to manage and co-ordinate road safety activities in a given country. For example, some have had success with the setting up of a national road safety co-ordinating committee, while others have one ministry (and often one senior person such as a minister) given sole responsibility for national road safety matters. A set of guidelines has been prepared by the Asian Development Bank (ADB), primarily for the Asia Pacific Region, which can help to improve road safety through co-ordinated action (ADB 1998).

Involving civil society

The importance of involving civil society in reducing accidents has been recognized in recent years. This has involved road administrations consulting with

local road users directly to determine their particular safety concerns, and also seeking their input into proposed improvements. An initiative by the World Bank called the ‘Global Road Safety Partnership’ (GRSP) began in 1999 with the aim of supporting and promoting strategic examples of partnerships for community development. GRSP’s function is to bring together government, businesses and civil society to improve road safety, and ensure that implementation is financially and institutionally sustainable. GRSP is at present focusing on 10 countries in Africa, Asia, Eastern Europe and Latin America.

The setting of targets is a well-established management strategy and, when applied to accident reductions, has proved very effective in many industrialized countries. A national target needs to be disaggregated so that all those with safety responsibilities, including road and traffic engineers, are given their own specific and realistic targets within this, with adequate funding related directly to the targets. It is also recommended that the authorities produce an annual road safety plan, which states clearly the current accident reduction target and how it is to be achieved. These plans should be published to enable public scrutiny, and to record the effectiveness of the road administration in improving road safety.

Plans and targets

Road safety organizations should be established on a full-time basis and be capable of

Organizational requirements

- diagnosing the road accident problem;
- drawing up an integrated plan of action, including the setting of targets;
- co-ordinating the work of all organizations involved;
- procuring funds and resources;
- producing design guides;
- designing and implementing improvements;
- monitoring implementation and evaluating measures;
- feeding back information from the evaluations and amending the action plan as necessary.

4.5.2 Road accident information systems

Diagnosis of the road accident problem is a key activity. This requires establishing a reliable database for the road network. The most important source of data for this activity is the police road accident report. Individual road administrations need to analyse their data to identify hazardous locations and the nature of the problems, to choose appropriate countermeasures and assess their effectiveness. Computer-based systems enable this to be done with increased efficiency and accuracy. Box 4.2 gives an example of a road accident information system.

4.6 Improving road safety

4.6.1 Engineering and planning

4.6.1.1 Mechanisms

Despite the fact that human error is usually the chief causal factor in road accidents, there is considerable scope for designing engineering and planning improvements to

Box 4.2 Microcomputer Accident Analysis Package

The Microcomputer Accident Analysis Package (MAAP) has been developed by Transport Research Laboratory (TRL), through collaboration with the traffic police in many countries. MAAP enables users to

- obtain good data for diagnosis, planning and research purposes;
- set up low-cost engineering improvement schemes, and to evaluate their own particular implementations of these.

The package consists of a set of software programs for data entry and analysis. These rely on accurate data from a police report. A basic standard structure is recommended for the report form, but details can be varied to adapt to local needs. MAAP has its own in-built map-based graphics enabling much analysis to be carried out directly using on-screen maps.

MAAP is now in use in over 50 countries, with 9 countries using it to provide their national database. MAAP has been configured to operate in several languages, including Arabic, Chinese, French and Spanish. The software has been developed gradually and improved over the years. For example, a Windows version is now available, which incorporates the latest techniques in data capture and analysis, and provides a range of tools for analysing problems. The latest version of MAAP is also used by a number of police and local authorities in the United Kingdom.

Source: TRL Ltd.

influence road-user behaviour so that errors are less likely to occur. The driving environment can also be made more ‘forgiving’ through relatively low-cost engineering techniques. There has been a greater emphasis on engineering and planning countermeasures over the past two decades both in Europe and North America. Engineering and planning can improve road safety through two distinct mechanisms:

- *Accident prevention* – resulting from good standards of design and planning for *new* road schemes and related development.
- *Accident reduction* – resulting from remedial measures applied to problems identified on the *existing* road network.

4.6.1.2 Accident prevention

Road design standards

There has been very little research in developing or emerging countries into the relationships between road design standards and accident rates. As a result, many countries have just adopted standards from developed countries or have modified such standards without evaluating the consequences. Often the traffic mix and road usage is very different in a developing country from that encountered in more

Box 4.3 Selected engineering and planning measures for accident prevention

Road hierarchy

- The categorization of roads in a network into a clearly defined road hierarchy – in particular, to separate roads whose primary function is movement from those which are primarily for access.
- Assign clear and unambiguous priority to each intersection so that traffic on the more important road is given precedence.

Land use

- Traffic and safety implications of all development proposals should be examined thoroughly before approval is given.
- Land uses should be distributed to minimize road traffic and pedestrian conflicts.
- The need for travel by motor vehicles should be minimized by locating shops and schools within walking distance of houses.

Access control

- On new arterial and collector roads, direct access to frontages should only be permitted in exceptional circumstances.
- The number of direct accesses onto main roads should be minimized, and service roads should be used to combine traffic in a single junction with the main road.
- No access should be permitted at potentially dangerous locations, such as intersections or bends with limited visibility.
- Any road should intersect only with roads of the same class (hierarchy), or with roads of no more than the next higher or lower class.

industrialized countries. To address this problem, a different approach to the geometric design of roads may be required in developing countries, especially for low-volume roads. Studies of the relationships between geometric design and road accidents in Kenya and Jamaica and research in Chile and India indicated, not unexpectedly, that apart from traffic flow, junctions per kilometre was the most significant factor related to accidents, followed by horizontal and vertical curvature. More research is required before optimum standards can be determined for all developing countries. Geometric design is discussed elsewhere in this book, but some tools available to the engineer and planner for the prevention of road accidents are summarized in Box 4.3. Some examples of good practice are illustrated in Figure 4.4.

The requirement for road safety audits has been adopted in a number of developing and emerging countries in recent years (IHT 1990a). These require that sound road safety principles are considered and checked by an independent team of safety specialists at the road planning, design and construction stages. This prevents road

Safety audits

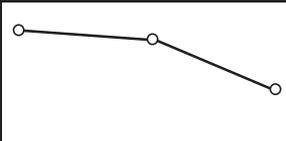
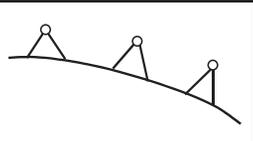
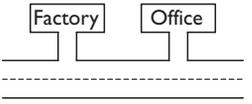
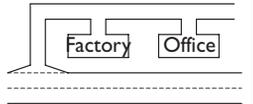
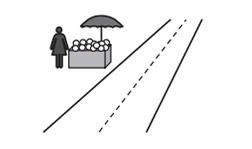
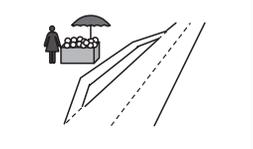
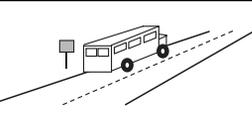
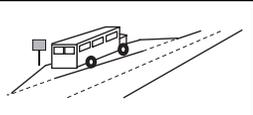
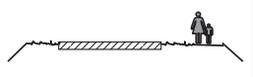
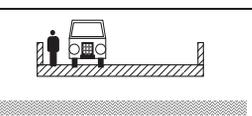
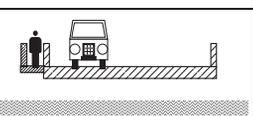
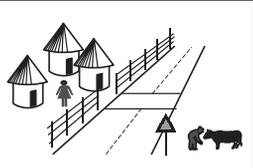
	<i>Undesirable</i>	<i>Desirable</i>	<i>Principle applied</i>
Route location			Major routes should by-pass towns and villages
Road geometry	(i)  (ii) 		Gently curving roads have lowest accident rates
Roadside access			Prohibit direct frontal access to major routes and use service roads
			Use lay-bys or widened shoulders to allow villagers to sell local produce
			Use lay-bys for buses and taxis to avoid restriction and improve visibility
Segregate motorized and non-motorized traffic			Seal shoulder and provide rumble divider when pedestrian or animal traffic is significant
			Construct projected footway for pedestrians and animals on bridges
			Fence through villages and provide pedestrian crossings

Figure 4.4 Examples of effects of engineering design on road safety.

Adapted from: TRRL Overseas Unit (1988).

safety being considered as an ‘after-thought’ when a new road is opened. Specific objectives of audits are

- to minimize the risk of accidents on a new road, and to minimize the severity of any accidents that do occur;
- to minimize the risk of accidents occurring on adjacent roads, and to avoid creating accidents elsewhere on the network;

- to recognize the importance of safety in road design to meet the needs and perceptions of all types of road user, and to achieve a balance where these needs may conflict;
- to reduce the long-term costs of road projects, bearing in mind that unsafe designs may be expensive or even impossible to correct at a later stage; and
- to improve the awareness of safe design practices by all involved in the planning, design, construction and maintenance of roads.

The roles and responsibilities for safety audits are shown in Box 4.4.

Box 4.4 Road safety audit roles and responsibilities

Commissioning audit (responsibility: client)

- Identify and prioritize projects requiring audit at different stages
- Specify terms of reference
- Approve audit team.

Initiating audit and provision of brief (responsibility: project manager/design engineer)

- Select audit team
- Provide terms of reference, any previous audits and exception reports
- Provide background information covering general description and objectives, plans, departures from standards, traffic and accident records.

Undertaking audit (responsibility: audit team)

- Assess information provided in brief
- Inspect project on site
- Formulate safety implications, potential problems and recommendations
- Assess using a check-list of features
- Present audit report
- Establish and maintain a documentation, storage and retrieval system for all audited projects.

Responding to audit (responsibility: project manager/design engineer)

- Consider audit report
- Agree changes in design to meet concerns and re-submit for re-audit
- Make exception reports on issues where no action is proposed to meet concerns
- Collate audit information.

Finalizing action (responsibility: client)

- Finalize decisions on changes to be made
- Refer any re-design to audit team
- Complete necessary documentation to demonstrate that audit recommendations have been considered and acted upon, with copies as appropriate to project manager/design engineer.

Guidance Unfortunately, in many cases, roads are still being built or upgraded with little consideration given to road safety and, as a result, hazards are still being created. In the past, it has been difficult to acquire information about appropriate techniques and standards. To address this, TRL has published *Towards Safer Roads in Developing Countries* (TRRL 1994), which is a road safety guide for planners and engineers. This guide is designed to be a first point of reference on road safety issues for road engineers.

4.6.1.3 Accident reduction

Low-cost improvements The approaches used by developed countries for accident reduction could have considerable potential for developing countries. In particular, it is recommended that countries with limited resources should place initial emphasis on introducing low-cost improvement schemes at hazardous locations. Such schemes have proved very effective in industrialized countries. The introduction of such schemes in the United Kingdom has produced first year rates of return estimated to range from 65 to 950 per cent. A few developing countries have begun to introduce such schemes on a trial basis. Preliminary findings suggest low-cost measures, such as road signs and markings, are less likely to be successful in countries with relatively low levels of road-user discipline. For example, a study of the effects of introducing stop lines and lane lines at junctions, and no overtaking lines at bends in Pakistan, indicated no improvements in driver behaviour apart from a small reduction in overtaking violations from 19 to 14 per cent. On the other hand, preliminary results from Papua New Guinea indicate that the introduction of roundabouts at uncontrolled major/minor junctions has halved the average injury accident rate (Hills *et al.* 1990).

Guidance Whereas industrialized countries have benefited considerably from improvements in engineering approaches to road safety, developing countries have been slower to adopt these. Several guides on the accident investigation and prevention process for engineers have been published and include, for example, Baguley (1995), Barker and Baguley (2001), IHT (1990b) and ROSPA (1992).

Applications There are some fields of engineering where many design standards from industrialized countries could be applied directly. One such application is in the area of street lighting, and a developing country manual has been published (ILE 1990). The manual predicts night-time accident savings of over 30 per cent for road lighting improvements in third world countries, although the costs of the improvements are relatively high compared with other measures. Another example is the application of 'traffic calming' methods that have proved to be so successful in developed countries (Lovell *et al.* 1994).

Counter measures Strategies for undertaking road accident countermeasures are described in Box 4.5.

4.6.2 Vehicle safety

Vehicle design and condition Improvements in vehicle design, occupant protection and vehicle maintenance have made a significant contribution to accident reduction in industrialized countries. In developing countries, however, the safety design of vehicles sometimes lags behind, particularly when vehicles are manufactured locally or assembled or imported second hand. Similarly, vehicle condition is likely to be poor when it is difficult to

Box 4.5 Countermeasures

Strategies

There are four basic strategies for accident reduction available to engineers through the use of countermeasures. These are

- Single site (blackspot programmes) – the treatment of specific types of accident at single location.
- Mass action plans – the application of known remedies to locations with a common accident problem.
- Route action plans – the application of known remedies along a route with a high accident rate.
- Area-wide schemes – the application of various treatments over a wide area of town/city, such as including traffic management and traffic calming (speed reducing devices).

Blackspot treatment is likely to be the most effective and straightforward in countries with no prior experience of accident remedial work. Many road administrations in industrialized countries started with this approach, and only later moved on to mass and route action plans as experience built up.

Data analysis

All of these strategies rely on the availability of data describing accidents and their locations to identify where accidents occur and what are the common features, which contribute to them. After the identification of accident blackspots, subsequent stages involve the analysis of data from each accident to establish the nature of the problem at the particular site. Accident analysis data and techniques, including both ‘stick diagram’ and ‘collision diagram’ analysis, are discussed fully in *Towards Safer Roads in Developing Countries* (TRRL 1994). Having identified common features and contributory factors, the next stage at each blackspot is to develop appropriate countermeasures. These commonly relate to

- Removing the conflict
- Improving visibility
- Reducing speeds.

Monitoring

Finally, it is important to monitor the impact of the ‘improvement’ over a period of time (often 2–3 years) and evaluate its effectiveness.

obtain spare parts. Overloading of goods and passenger vehicles is another factor that commonly contributes to high accident severity and casualty rates.

The control of overloading passenger-carrying vehicles, combined with improvements in the design of such vehicles, has some potential for accident and casualty reduction in many countries. For example, in Papua New Guinea, it is common for passengers to be transported in open pick-ups and, perhaps not surprisingly,

Passenger
overloading

an exceptionally high proportion (45 per cent) of the road accident casualties come from such vehicles. To help Papua New Guinea deal with this problem, a simple, robust protective cage was designed by TRL to protect occupants. Rollover trials on a test track demonstrated that the cage provided improved protection, with the potential for reducing the severity of injuries and to prevent vehicle rollover in certain conditions.

Guidance

From Table 4.3 it is clear that the police in some developing countries have blamed a relatively high proportion of accidents (up to 17 per cent) on vehicle defects. Although many of these countries may have inadequate controls to ensure minimum safe standards of vehicle condition, it would be more appropriate to start by introducing low-cost random roadside checks using simple equipment rather than expensive networks of vehicle testing centres with sophisticated technology. The benefits to individual road users of improved vehicle design and of occupants wearing seat belts and helmets are likely to be much the same from one country to another. Thus, the general adoption of both primary and secondary vehicle safety measures is to be encouraged. However, the total benefit of such measures to a developing country as a whole will depend on the characteristics of its accident and casualty problem and, in some cases, on the degree of road users' compliance with traffic legislation. For example, seat belt laws would lead to only small casualty savings if few casualties come from cars, or if most drivers and passengers ignore the law. Indeed, studies in India (Mohan 1992) have indicated that, if helmet-wearing law were properly enforced, then the death rate among motor cyclists due to head injuries would be reduced by 30–40 per cent.

4.6.3 Education and training

4.6.3.1 Road safety education

Children's education

It is important for road users to be educated about road safety from as young an age as possible, with simple messages about the dangers of roads and traffic. In industrialized countries, a number of approaches have been tried, both through school systems and through parents, and most children receive some advice. However, in developing countries, where the child pedestrian accident problem is generally more serious, a study of children's crossing knowledge (Downing and Sayer 1982) indicated that children were less likely to receive advice than in the United Kingdom from members of their family, teachers or the police. There is clearly a need to improve road safety education. However, since some countries have low school attendance figures, it is important that education through community programmes is considered as well as at school.

Teaching methods

A number of studies in Europe have evaluated teaching environments for children crossing roads. Overall, the results demonstrated the importance of training on real roads. The need for frequent supervised practice on local roads close to where children live should be emphasized. Road safety education programmes should be graded and developmental. Teachers need guidelines on what and how to teach. To meet these requirements, many countries have produced syllabus documents and teacher guides (Leburu 1990; Sayer *et al.* 1997). However, the transferability of industrialized country solutions to developing countries is uncertain in these areas.

4.6.3.2 Driver training and testing

In developing and emerging countries, the problems of poor driver behaviour and knowledge are likely to be due, to some extent, to inadequacies in driver training and testing. Professional driving instruction tends to be inadequate because

Driver
instruction

- driving instructors are not properly tested or monitored;
- there are no driving or instruction manuals;
- driving test standards and requirements are inadequate.

Consequently, there is likely to be considerable scope for raising driving standards by improving driver training and testing. Many countries also need to improve the licensing, training, testing and monitoring of instructors to ensure that appropriate driving standards are taught.

In training systems where learner drivers are free to choose how they learn, it is important that driving tests demand a high standard of driving, especially for the practical ‘on-the-road’ assessment. More difficult tests should encourage learners to purchase lessons from professional instructors.

Driving tests

Truck drivers tend to have a greater involvement in accidents than in industrialized countries, and inadequate training plays some part in this. A driving guide specifically for truck drivers has been produced by TRL in collaboration with the United Nations Economic Commission for Africa (TRRL 1990). A study demonstrated that reading sections of the guide helped drivers improve their scores on knowledge tests by up to 25 per cent on some topics.

Truck drivers

There has been little research on the effectiveness of improved driver training in developing and emerging countries. Accident savings as a direct result of training are, of course, very difficult to prove. A study of a retraining course for bus drivers in Pakistan (Downing 1988) failed to demonstrate any accident savings, although there was evidence of an improvement in knowledge test scores (13 per cent on average) and a reduction in driving test errors (67 per cent on average). It was also found that the training had no effect on the drivers’ behaviour when they were observed unobtrusively, and they clearly returned to their old habits when driving in normal conditions. Therefore, to bring about a general improvement in driver behaviour, it will usually be necessary to ensure that drivers are sufficiently motivated. Training courses will probably need to be integrated with publicity campaigns, incentive schemes and enforcement.

Effectiveness

4.6.4 Enforcement

Many studies of the effectiveness of enforcement have demonstrated that a conspicuous police presence leads to improvements in driver behaviour in the vicinity of the police, but the evidence for accident reductions is less convincing. In developing and emerging countries, traffic police are generally poorly trained and equipped and, often, they are not mobile, being stationed at intersections. Traffic police operating under such conditions are likely to find it difficult to influence traffic violations, and this was confirmed by a study of the effects of police presence in Pakistan. However, the study also showed that improved training and deployment of traffic police

Effectiveness
of policing

produced large reductions in traffic violations. Following the introduction of road patrols on inter-city roads, a 6 per cent reduction in accidents was achieved in Pakistan, and a similar scheme in Egypt produced accident reductions of almost 50 per cent. Therefore, it would appear that improvements in traffic policing, as indicated, have considerable potential for both improving driver behaviour and reducing accidents, provided that the police's capability to enforce traffic violations is enhanced.

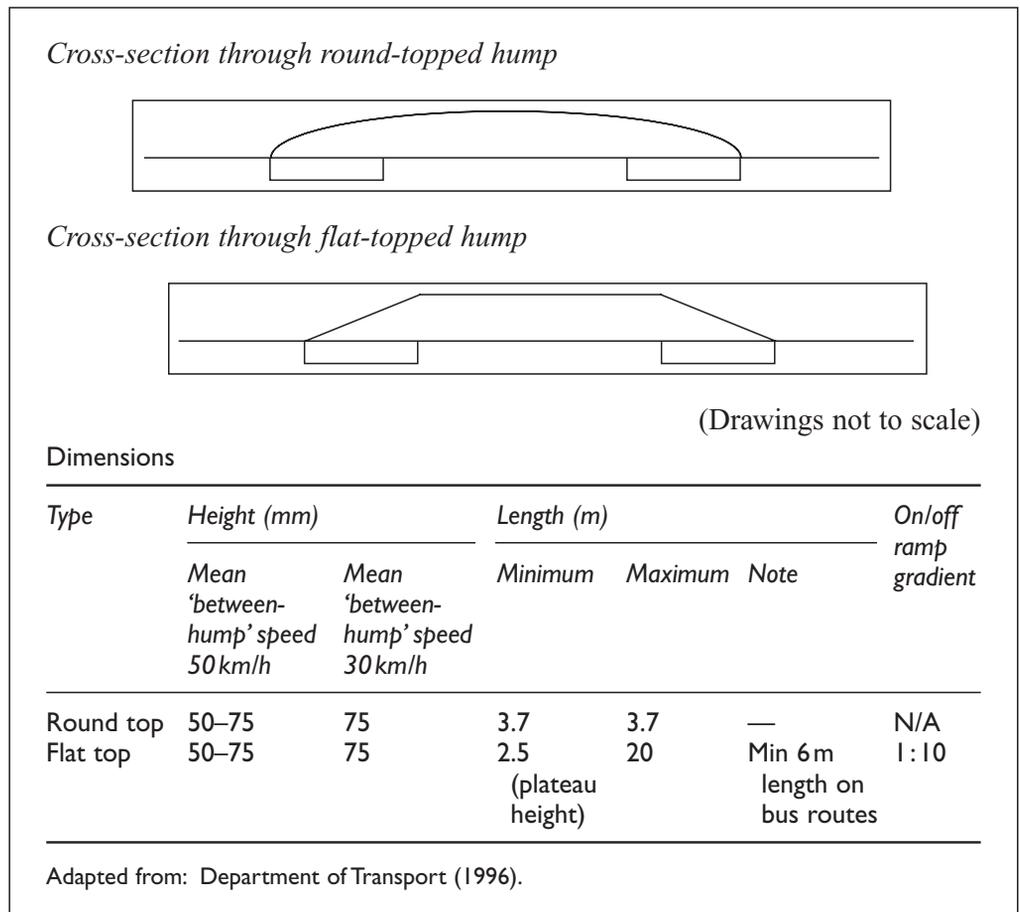
Speed

Much attention has recently been devoted to the management of vehicle speeds since the finding (Taylor *et al.* 2000) that each mile per hour (1.6 km/h) reduction in average speed is expected to cut accident frequency by 5 per cent, with slightly different percentages depending on road type. The United Kingdom Government is currently tackling this by ensuring speed limits are set appropriately, signing is improved, speeding penalties reviewed, more traffic calming devices introduced, and greater use made of speed cameras. An appropriate design for road humps to limit vehicle speed is shown in Box 4.6.

Publicity campaigns

Research in industrialized countries suggests that changes in traffic police operations need to be well advertised to ensure maximum effect on road-user behaviour. This finding is likely to be universal, so it is therefore important that developing and

Box 4.6 Type and dimensions of road hump



emerging countries integrate changes in enforcement tactics with appropriate publicity campaigns. In many countries, it is likely that such improvements will need to be accompanied by modification to both the traffic legislation and the ways of dealing with offenders. For example, Victoria in Australia is widely reported as having the greatest success in reducing drink driving. This has been reduced to about one-third of its previous level over a 6-year period, and fatalities involving alcohol have been halved. This has been achieved through a combination of hard-hitting advertising campaigns and strict enforcement using a fleet of 'booze buses', as well as mobile breath test units in patrol cars.

4.6.5 Pilot trials

The need to reflect local circumstances and the lack of data on the benefits of road safety measures suggest that, if possible, improvements are introduced on a pilot basis and evaluated before being implemented nation-wide.

4.7 Priorities for action

Road accidents in developing and emerging countries are

- a serious problem in terms of fatality rates, which are at least an order of magnitude higher than those in industrialized countries;
- an important cause of death and injury;
- a considerable waste of scarce financial (and other) resources, typically costing between 1 and 4 per cent of a country's gross national income per annum.

In Europe and North America the road accident situation is generally improving, but many developing and emerging countries face a worsening situation. Whereas high-income countries have had over half a century to learn to cope with the problems of ever-increasing motorization, the less wealthy nations have had less time and, for many, the pace of change has been much greater.

It is important that there is a clear understanding of the road accident problem and the likely effectiveness of road safety improvements. A priority for countries is, therefore, to have an appropriate accident information system that can be used to identify accident patterns, the factors involved in road accidents and the location of hazardous sites. Procedures for costing road accidents need to be set up to enable a budget for a 5-year action programme to be determined. This will also ensure that the best use is made of any investment, and that the most appropriate improvements are introduced in terms of the benefits that they will generate in relation to the cost of their implementation. Other basic requirements for the lower- and middle-income countries are likely to include the following:

- Adopt a scientific, quantitative basis for road safety policy, and integrate this with national transport policy; establish research centres, establish data systems.
- Create and strengthen institutions so that they can prepare and deliver multi-sectoral road safety plans and actions; form road safety councils, train safety

Summary of
problem

Priorities

teams and establish realistic targets; adopt partnership approaches involving government, civil society and business.

- Press for long-term land use and transport policies to reduce the more dangerous modes and mixes of traffic.
- Plan well-defined road hierarchies; use appropriate and consistent geometric design standards; reduce the unexpected on roads through safety audits and low-cost road improvements; introduce low-speed/pedestrian priority areas in cities.
- Target vulnerable, young and elderly road users and those largely responsible for accidents through education and training; legislate against and control drink (and drug) driving and speeding; improve targeted enforcement.
- Enforce the use of helmets for two-wheelers and seat belts for car and heavier vehicle occupants; improve general vehicle condition as well as other measures for public and parastatal freight and bus fleets.
- Improve emergency medical services for those who survive the initial impact of the crash but die within the next few hours (often two-thirds).
- With all measures, adapt, experiment and evaluate; share the lessons learned on a national and international basis; improve plans and scale up successes.

References

- ADB (1998) *Road Safety Guidelines for the Asian and Pacific Region*, Manila: Asian Development Bank.
- Baguley, C.J. (1995) 'Interim guide on identifying, prioritising and treating hazardous locations on roads in Malaysia', *JKR 2078-022-95*, Kuala Lumpur: Jabatan Kaya Raya, Ministry of Public Works.
- Barker, J. and Baguley, C.J. (2001) *A Road Safety Good Practice Guide*, First edn, London: Department for Transport, Local Government and the Regions.
- Department of Transport (1996) 'Highways (road humps) regulations 1996', *Traffic Advisory Leaflet 7/96*, London: Department of Transport.
- Downing, A.J. (1988) 'The effectiveness of a retraining programme for bus drivers in Pakistan', in Rotengatter, J.A. and Bruin, D.E. (eds), *Road User Behaviour: Theory and Research*, Assen/Maastricht: Van Gorcum, 682–689.
- Downing, A.J. and Sayer, I.A. (1982) 'A preliminary study of children's road crossing knowledge in three developing countries', *Supplementary Report 771*, Crowthorne: Transport and Road Research Laboratory.
- Harral, C. and Faiz, A. (1988) 'Road deterioration in developing countries: causes and remedies', in *A World Bank Policy Study*, Washington, DC: World Bank.
- Hills, B.L., Thompson, N. and Kila, K. (1990) 'Accident blackspot research in Papua New Guinea', *Proceedings of the Fourth Annual Conference on Road Engineering, Jakarta, November 19–21*.
- Hills, B.L., Thompson, N., Johnson, F. and Posnau, C.D. (1996) 'A roadside and hospital survey of drinking and driving in Port Moresby, Papua New Guinea', *Research Report 211*, Crowthorne: Transport Research Laboratory.
- IHT (1990a) *Guidelines for the Safety Audit of Highways*, London: Institution of Highways and Transportation.
- IHT (1990b) *Highway Safety Guidelines: Accident Reduction and Prevention, International Edition*, London: The Institution of Highways and Transportation.
- ILE (1990) *A Manual of Road Lighting in Developing Countries*, Rugby: Institute of Lighting Engineers.

- Jacobs, G.D., Downing, A.J. and Sayer, I.A. (1981) 'A preliminary study of road user behaviour in developing countries', *Supplementary Report 646*, Crowthorne: Transport and Road Research Laboratory.
- Jacobs, G.D., Aeron-Thomas, A. and Astrop, A. (2000) 'Estimating global road fatalities', *Research Report 445*, Crowthorne: Transport Research Laboratory.
- Leburu, F.M. (1990) 'Training the children: traffic education in Botswana', *Proceedings of the International Conference for Road Safety and Accidents in Developing Countries*, Cairo: The Academy of Scientific Research and Technology.
- Lovell, A., Baker, M., Bennett, J., Bulpitt, M., Clayton, J., Endersby, R., Martin, J., Primmer, P., Talbot, M. and Thame, C. (1994) *Traffic Calming in Practice*, London: Landor Publishing on behalf of County Surveyors Society and Department of Transport.
- Mohan, D. (1992) 'Safer two wheelers and helmets', *International workshop on Prevention and Control of Traffic Accidents and Injuries, New Delhi*, 143–148.
- Murray, C.J.L. and Lopez, A.D. (1996) *The Global Burden of Disease*, Cambridge MA: Harvard University Press on behalf of World Health Organization and World Bank.
- ROSPA (1992) *Road Safety Engineering Manual*, Birmingham: Royal Society for the Prevention of Accidents.
- Sabey, B.E. and Staughton, G.G. (1975) 'Interacting roles of road environment, vehicle and road-user in accidents', in IAATM, *Fifth International Conference, London, September 1975*, London: International Association for Accident and Traffic Medicine.
- Sayer, I.A., Palmer, C.J., Murray, G. and Guy, J. (1997) 'Improving road safety education in developing countries', *Research Report 265*, Crowthorne: Transport Research Laboratory.
- Taylor, M., Lynam, D. and Baruva, A. (2000) 'The effect of drivers' speed on the frequency of road accidents', *Research Report 421*, Crowthorne: Transport Research Laboratory.
- TRL (1995) 'Costing road accidents in developing countries', *Overseas Road Note 10*, Crowthorne: Transport Research Laboratory.
- TRRL Overseas Unit (1988) 'A guide to geometric design', *Overseas Road Note 6*, Crowthorne: Transport and Road Research Laboratory.
- TRRL (1990) *A Guide for Drivers of Heavy Goods Vehicles*, Crowthorne: Transport and Road Research Laboratory.
- TRRL (1994) *Towards Safer Roads in Developing Countries: A Guide for Planners and Engineers*, Crowthorne: Transport and Road Research Laboratory.
- WHO (1999) *World Health Report: Making a Difference*, Geneva: World Health Organization.

Roads and the environment

William V. Kennedy

5.1 Introduction

Constraints
and
opportunities

Road engineers have always recognized the importance of the physical environment. Their principal concern, however, has been whether natural features pose engineering constraints or opportunities for road construction. Roads, however, can have an important impact on the environment as well as being impacted by it.

Environmental
effects

Probably the most obvious environmental effect of roads is the physical one on the landscape. The narrow lanes and roads that existed before the car generally conformed to the natural landscape. Fast moving cars and trucks, however, need a different kind of road. If a vehicle is to drive fast, it cannot be burdened by many sharp curves. If a road is to have fewer curves, it must automatically have more bridges, cuts and fills. These can dramatically affect the landscape, destroying farmland, forests and other natural areas, as well as human and wildlife habitats. Although the impact on the natural landscape is the most visible, it is not the only environmental effect of roads. The effects on air and water quality, as well as noise, cultural and social impacts are equally important.

Issues

This chapter begins by describing the environmental impacts most commonly associated with roads and the mitigation measures for reducing the magnitude of their effects. It is followed by a description of the requirements and procedures for environmental impact assessment (EIA), which has emerged as the primary environmental planning and design tool for ensuring that roads are planned and operated in an environmentally sound manner.

5.2 Environmental impacts and their mitigation

5.2.1 Water resources

Areas of
impact

Road construction can lead to changes in three types of water resources: surface water flows, ground water flows and water quality degradation.

Surface water
flows

The effect on surface water flows is evidenced through flooding, soil erosion and channel modification. By modifying the natural flow of surface water, stream channels are disrupted by artificial changes in channel depth, width or shape. Such changes are associated with stream relocation, channelization, confinement in a culvert, or disruption by bridges or abutments. Straightened channelized streams usually permit faster passage of water than in undisturbed streams. This may lead to erosion

or downstream flooding. Alternatively, water backed up at the constriction may cause bank overflow and upstream flooding.

Road construction can affect ground water flows through the compaction of sub-surface materials. The creation of a sound foundation for roads often requires extensive drainage of sub-foundation layers in places where the natural water table approaches the surface. These activities may disrupt the rate of direction of ground water flow much like a dam across a stream. This can, in turn, lead to erosion, deterioration of soil and vegetation, loss of water for drinking and agricultural use and impacts on fish and wildlife.

Ground water flows

Impact on water quality most often results from sedimentation from construction activities, as well as runoff and spills once roads are in operation. Sedimentation results from the erosion of surface areas and the transport of the material by wind or water to a new resting place. When eroded material reaches surface waters, it either settles to the bottom, altering stream bottom conditions, or it remains in suspension as a pollutant. The ecological consequences of sedimentation include a reduction in sunlight reaching green plants, physical damage to aquatic organisms and destruction of bottom habitat required by aquatic species. Extensive sedimentation can also reduce the water storage capacity of ponds and reservoirs, and increase the likelihood of flooding by reducing the storage capacity of streams.

Water quality

Road construction may also lead to degradation in ground water quality. Ground water is normally very pure. Removal of protective impervious surface layers, through road works, may expose underlying areas to contamination from pollutants carried by surface flow.

Ground water quality

Negative impacts on water resources can be mitigated in a number of ways. Often the most effective and basic of these is, simply, choosing an alignment that has the least amount of impact on the resources; that is a minimum number of water crossings; avoidance of ground water protection zones, wetlands, etc. Water speed reduction measures included in road design can also substantially reduce impacts. These include actions such as grass planting, as well as the provision of check dams and boulder clusters in channels, to slow water velocity and reduce erosion. The construction of settling basins to intercept and remove silts, pollutants and debris from road runoff water, before it is discharged to adjacent streams or rivers, can also be effective. Interceptor ditches can also be used to reduce the flow of pollutants. A comparison of a number of these measures, in terms of their cost and effectiveness, is provided in Figure 5.1.

Mitigation measures

5.2.2 *Flora and fauna*

The effects of roads on flora and fauna, that is the plants and animals that make up the surrounding natural environment, can be of both a direct and indirect nature. Direct impacts can result from both wildlife mortality (animals killed attempting to cross the road) as well as the loss of wildlife habitat. In addition, the severance of habitats, brought about through the construction of a road in a remote, natural area can often be as damaging as the actual amount of habitat loss. A new road can isolate animal and plant populations living on either side of the right-of-way, leading in some cases to their extinction. Severance can also restrict the access of some animals to their usual areas of reproduction or feeding and, thereby, also lead to their demise.

Direct impacts

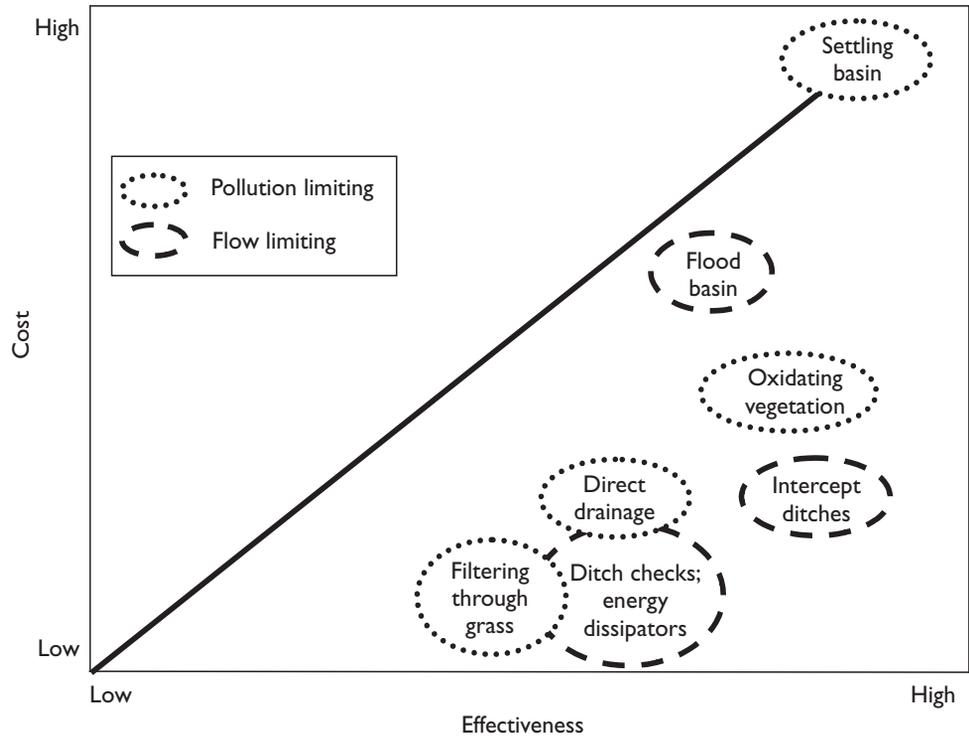


Figure 5.1 Comparison of water impact mitigation measures.

Source: R. Scott Hanna, Acres International.

Wetlands

Road construction in wetland areas can be particularly damaging to flora and fauna, and to water resources. Wetlands are important for the water cycle because they serve as natural flood control reservoirs and contribute to the base flow of streams. As habitat, they serve as the breeding, nursery, feeding and resting areas for many wild animals.

Indirect impacts

The indirect impacts to flora and fauna are related mainly to the increased human access to wildlife habitat and sensitive ecosystems that are facilitated by road provision. Other indirect impacts relate to the ecological disequilibrium which can, for example, come about as the result of wetland destruction, as described earlier. Contamination through road traffic accidents can also be significant.

Mitigation measures

The best way to mitigate against these impacts is to avoid them altogether by locating roads away from areas with sensitive flora and fauna populations. Where this is impossible or too costly to do, there are a number of specific steps that can be taken in road design and operation to reduce impacts. These include

- Engineering road cross-section designs to maintain habitat and avoid vehicle–wildlife collisions through narrower clearing of right-of-way, lower vertical alignment and improved sight lines for drivers.
- Provision of wildlife underpasses or overpasses to facilitate animal movement, and fencing to restrict movement and direct wildlife to designated crossing points.

In some cases, such as with the destruction of wetlands, the best mitigation can be obtained by creating new wetlands away from the project area to compensate for the

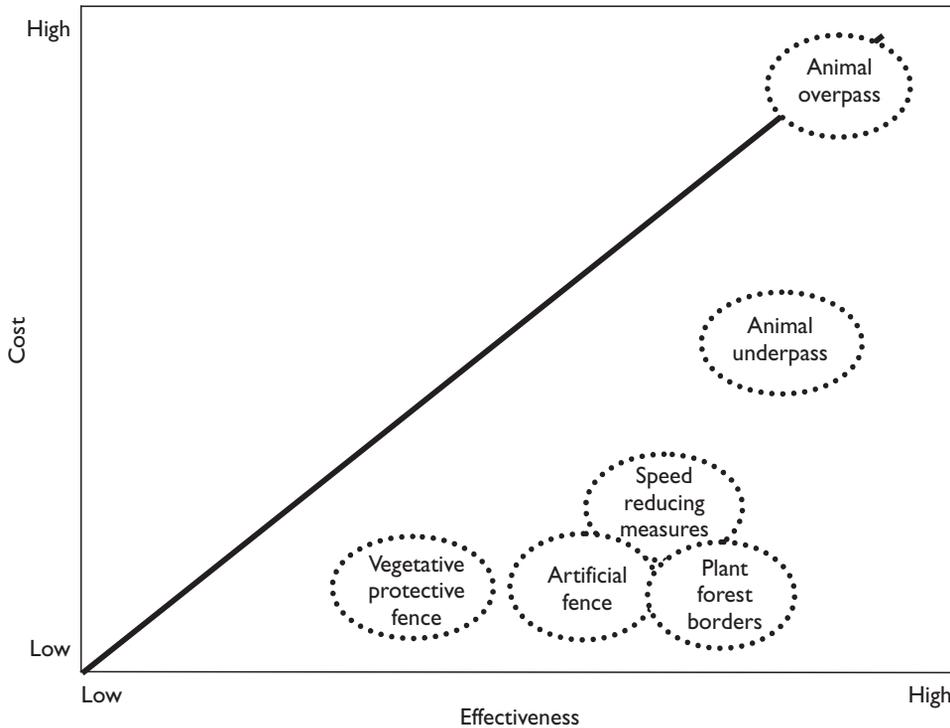


Figure 5.2 Comparison of flora and fauna protection mitigation measures.

Source: R. Scott Hanna, Acres International.

loss of the original area. A comparison of a number of these measures in terms of their cost and effectiveness is provided in Figure 5.2.

5.2.3 Soils

Soil is important as a medium for ecological and agricultural activities. However, it is also important for use as the foundation for roads and structures, whether obtained from within the road line, or imported from borrow pits. As such, it provides a good example of the kind of 'environmental factor' that both affects, and can be effected by, road construction.

Importance

The most direct impact on soils is erosion. Erosion can result from many causes, which are the consequence of the interaction between soil structures, climatic conditions and water resources. The consequences can extend beyond the immediate vicinity of the road where, for example, a reduction in slope stability and resultant landslides can have an immediate impact. There can also be far-reaching effects on streams, rivers and dams that are some distance away.

Erosion

Further impacts on soils include contamination, resulting from the disposal of excavated soils contaminated by past vehicle emissions. Contamination can also result from spills, runoff and the transportation of hazardous products during road construction and subsequent maintenance. The removal of agricultural soils, when roads are constructed in farming areas, represents a further impact.

Contamination

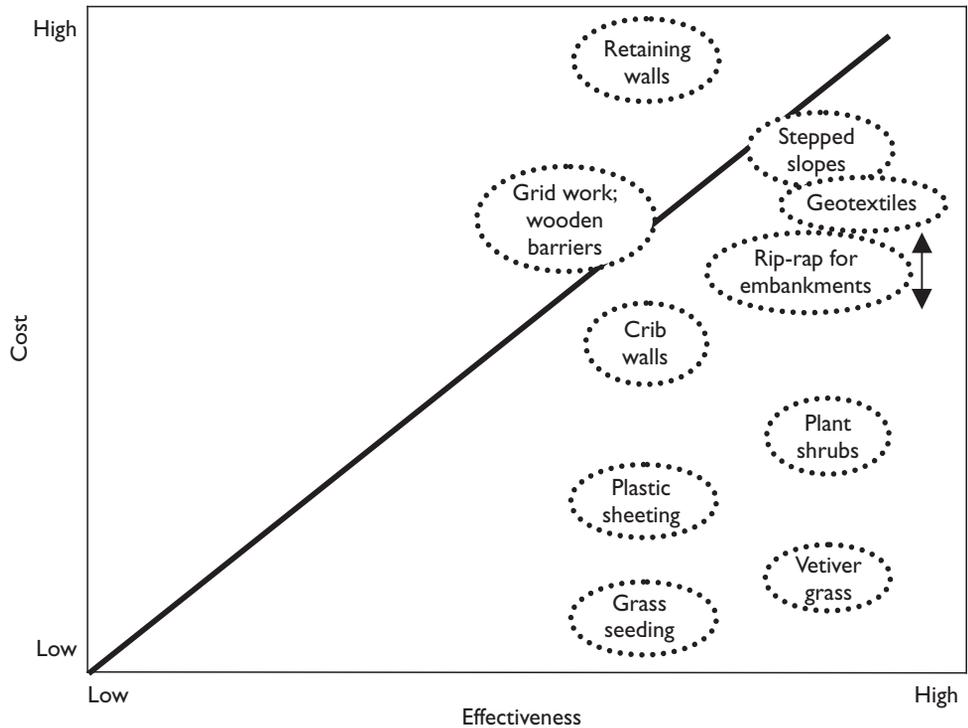


Figure 5.3 Comparison of erosion mitigation measures.

Source: R. Scott Hanna, Acres International.

Mitigation measures

Measures for avoiding or minimizing soil erosion include

- Minimizing the size of cleared areas
- Replanting cleared areas and slopes
- Storing and re-using topsoil through a number of engineering measures, such as interception ditches, contoured slopes, rip-rap, retaining walls, etc.

In order to avoid damaging soil productivity, road planners should, wherever possible, avoid alignments through highly productive agricultural lands. Where this is not possible, valuable topsoil should be stored and re-used. Quarries and borrow pits should be transformed into lakes and roadside amenities. A comparison of mitigation measures in terms of cost and effectiveness is provided in Figure 5.3.

5.2.4 Air quality

Pollutants

The construction, operation and maintenance of a road may substantially change air quality by introducing significant amounts of pollutants into the air.

Particulates

Particulates from cars, trucks and buses include

- Carbon particles
- Lead compounds and motor oil droplets from exhaust systems
- Oil from engine crankcase exhaust
- Particulate rubber from tyres
- Asbestos from brake and clutch linings.

Particulates cause inflammation of the lungs and increase incidents of pneumonia. In addition, they may carry irritants and carcinogenic material into the respiratory system and encourage the development of diseases, such as bronchitis and lung cancer.

Carbon monoxide (CO) is the most abundant pollutant in motor vehicle exhaust. Although inherently less noxious than other pollutants, it represents the greatest danger to health, since it is concentrated in areas of traffic congestion and tunnels. The danger to humans arises from a strong affinity for haemoglobin – the oxygen-carrying red cells of the blood. The known effects range from a headache to death, depending on the degree of concentration and duration of exposure.

The oxides of nitrogen emitted by motor vehicles are nitric oxide (NO) and nitrous oxide (N₂O). NO is a direct product of high temperature combustion. N₂O is a secondary product, resulting from chemical reactions in the atmosphere. Like CO, N₂O has an affinity for blood haemoglobin. It also forms acid in the lungs causing inflammation, cellular changes in lung tissue and the development of lesions.

The physiological effects of sulphur dioxide (SO₂) are well documented. It causes an irritation of various parts of the respiratory system, decreases mucus clearance rates, changes sensory perception and lessens pulmonary flow resistance. It can remain in the atmosphere for a number of weeks after being emitted.

Hydrocarbons (HC) are created by the incomplete combustion of fuel and its evaporation. They include a wide variety of organic chemical substances. They have multiple effects on health: toxic, irritant, carcinogenic or mutagenic.

Other pollutants include benzene, carbon dioxide (CO₂) and chlorofluorocarbons (CFCs). These primary pollutants can be transformed into secondary and tertiary pollutants, such as ozone and acid rain, through various chemical reactions linked to meteorological factors, air temperature and humidity.

Mitigating the negative effects of these pollutants is often beyond the competence or responsibility of the road engineer, since they are most directly connected with measures such as fuel technology and quality, as well as vehicle emissions standards, and inspection and maintenance requirements. Nonetheless, there are a number of actions that can be taken as part of the road planning and construction process to lessen the negative impacts on air quality. These include

- Routing traffic around residential areas and sensitive ecosystems.
- Use of vegetation to filter dust and particulate matter.
- Ensuring that construction equipment is in sound operation condition.
- Locating processing facilities, such as asphalt batch plants, away from residential areas.

5.2.5 Noise

Noise is often regarded as the most directly unpleasant impact of roads. While impacts on air and water quality and natural ecosystems are not always easily detected by the observer, noise impacts are readily apparent. The noise intensity of a single truck, for example, is 10,000 times greater than that of the natural noises in a deciduous forest environment. Consequently, this impact is increasingly being recognized as a source of environmental degradation. However, in developing and emerging countries its effects are often viewed with lower priority than economic,

Carbon
monoxide

Nitric and
nitrous oxide

Sulphur
dioxide

Hydrocarbons

Other
pollutants

Mitigation
measures

Impacts

Sources of noise	<p>health or other environmental impacts. Although the health effects of road noise have not been well documented, studies indicate that prolonged exposure to industrial noise levels may cause harm to human beings. Moreover, noise has also been shown to interfere significantly with wildlife nesting and movement habits.</p> <p>All transport modes are potential sources of noise, but the impacts from road traffic are the most widespread. The perceived impact is dependent on several factors:</p> <ul style="list-style-type: none"> • The type, number and speed of vehicles. • The characteristics of the interface between vehicle and the road surface. • Whether the environment is urban, suburban or rural. • The distance from the road to the receptor.
Mitigation measures	<p>Operation of heavy equipment during road construction is also a source of noise, though of only a temporary duration. The operation of loaders, tractors, graders and other equipment can result in noise levels higher than those produced by heavy traffic.</p> <p>The most effective noise mitigation measure is avoidance, such as by moving the road alignment or diverting traffic away from built-up areas. Motor vehicle noise can also be reduced at its source through improved vehicle construction and maintenance. As with the air pollution, however, these measures are beyond the scope of work of the road engineer. The most common form of engineering mitigation is the use of noise barriers. These are usually constructed from wood, metal or concrete and are known as barriers, fences or screens. They are often combined with planting schemes which, in themselves, do not offer much protection against noise, but are effective in softening the visual appearance of barriers. Other mitigation measures include the installation of double glazed windows and/or insulated building facades. During construction, workers should be equipped with hearing protection. A comparison of mitigation measures in terms of cost and effectiveness is provided in Figure 5.4.</p>
<p>5.2.6 Cultural heritage</p>	
Impacts	<p>The term ‘cultural heritage’ usually refers to cultural/historical monuments and archaeological sites. The construction of roads can have a direct impact on this heritage, and particularly on archaeological sites. Often their discovery is as a result of road construction. Physical damage to cultural sites can also be caused by quarry and borrow site works, as well as unregulated access to construction sites.</p>
Mitigation measures	<p>The main mitigation measure is to avoid alignments through known cultural sites and to realign the road if an important site is uncovered during construction. In addition, site management plans can also be developed, elements of which may include requirements to:</p> <ul style="list-style-type: none"> • Carry out mapping of archaeological sites • Develop contract clauses for construction • Excavate and relocate artefacts • Stabilize embankments, soils and rock containing artefacts

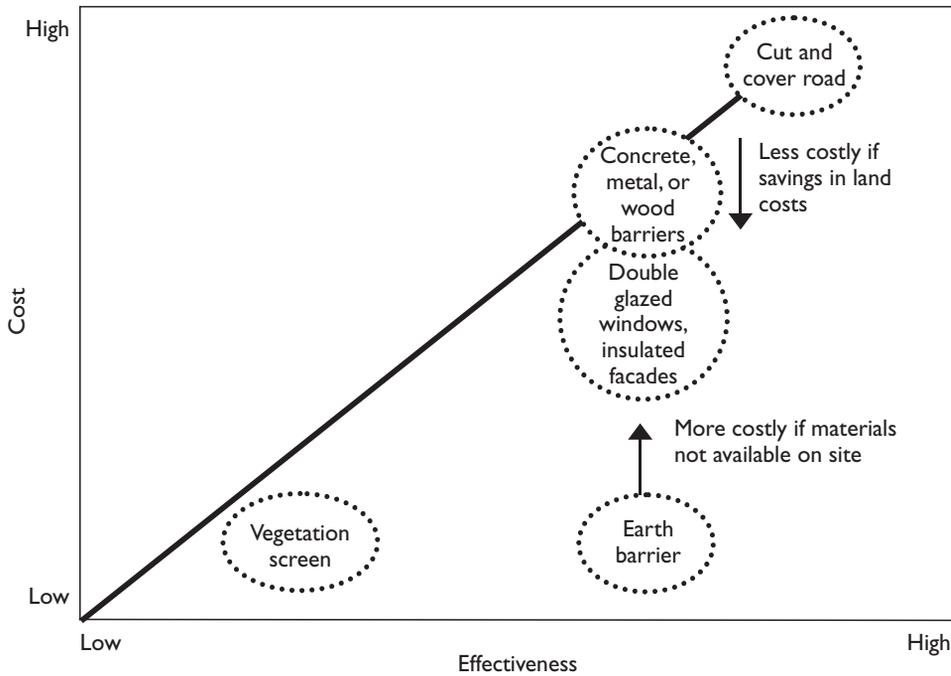


Figure 5.4 Comparison of noise mitigation measures.

Source: R. Scott Hanna, Acres International.

- Control groundwater levels
- Control flora and fauna
- Establish monitoring and evaluation systems.

5.2.7 Socio-economic issues

Roads are provided to bring socio-economic benefits through improved access, lower transport costs and better markets for goods and services. However, often the negative socio-economic impacts can be greater than the impacts on the natural environment. For example, bigger roads and increased traffic can make it difficult for local inhabitants to cross roads, reduce accessibility to local activities adjacent to the right-of-way, and disrupt the traditional patterns of everyday life and business. These 'severance' effects are difficult to quantify and are a frequent cause of community concern with road projects in populated areas. Other negative socio-economic impacts can result from roads which by-pass communities and bring about a reduction in business. Ironically, these same by-passes can result in positive environmental impacts to the same communities through a reduction in traffic noise, congestion and pollution.

Tourism and agriculture are two economic activities that can be both positively and negatively affected by road provision. While improved access may positively benefit tourism, increased activity could damage tourist attractions and disrupt farming patterns and connections between fields.

Impacts

Tourism and agriculture

Mitigation measures As with many environmental impacts, one of the best ways to mitigate negative socio-economic impacts is to locate new roads away from human settlements. Severance can be minimized by taking account of local movements at the road design stage, and by making provision for improved crossings or alternative access routes, or both. This can be achieved through the use of traffic signals, pedestrian underpasses and overpasses, and by the provision of service roads. In the final analysis, often the most effective means of addressing socio-economic impacts is through compensation. This may include resettlement and monetary compensation for persons or businesses directly affected, as well as through the provision of roadside improvements and community amenities.

5.3 Environmental impact assessment

5.3.1 Background

Historical context Since the concept was first introduced in the *US National Environmental Policy Act* of 1960, EIA has emerged as the dominant tool for ensuring that the environmental impacts of development activities are considered in decision-making. Over the last 40 years, more than 100 countries have introduced some form of EIA into their national planning processes. During this period, more environmental impact assessments have probably been prepared for transport activities, and particularly road projects, than for any other sector.

International agencies EIA procedures and processes have also been adopted by a number of inter-governmental organizations, such as the United Nations, the Organization for Economic Co-operation and Development (OECD) and the European Commission. Bilateral aid agencies and multilateral financial institutions have made EIAs part of the decision-making process for investments. This has been responsible for the bulk of the EIAs carried out on road projects in developing countries and, more recently, in the economies in transition of Eastern Europe and the former Soviet Union. The following sections are based on the policies and procedures of the World Bank, the regional development banks and the bilateral agencies.

5.3.2 Screening

Classification The EIA process begins normally with the ‘screening’ of investment projects into categories. A three- or four-level classification system is normally used for this – see, for example, the *Environmental Procedures* produced by the EBRD (2003). The categories serve to classify projects according to their potential environmental impacts. Classification into a certain category indicates the extent, magnitude and significance of environmental issues, and the corresponding level of assessment and review required to address them.

Screening categories The three main screening categories (sometimes referred to as ‘A’, ‘B’ and ‘C’) signify the potential of negative environmental impact and can be labelled generically as

- High impact – full environmental appraisal (EIA) required
- Moderate impact – limited or partial environmental appraisal required
- Low impact – no environmental appraisal required.

Road projects involving the design and construction of new motorways, or other primary roads, routed through ‘greenfield’ sites are screened in the ‘high impact’ category, requiring an EIA. Road upgrading projects, involving widening or the provision of additional lanes, are screened in the second ‘moderate impact’ category. These require a lower level of environmental appraisal, often called an ‘environmental analysis’. Road projects that are limited to rehabilitation or maintenance, such as the provision of new asphalt surface, are screened in the third category, and do not require an environmental appraisal.

5.3.3 Scoping

For road projects screened in the high impact category, the first step in preparing the EIA is termed ‘scoping’. This is the process for determining the range of issues and impacts to be considered in the EIA, and the way in which the public will be involved in the appraisal. It is usually carried out through a scoping meeting called by the project sponsor to which are invited representatives of the locally affected public, as well as representatives of environmental authorities, municipalities, other government departments and NGOs. The aim is to discuss and agree upon the ‘scope’ of the EIA.

5.3.4 Assessment methods and techniques

On the basis of the scoping exercise, the final terms of reference or table of contents for the EIA are established. A number of guidelines and manuals have been developed for identifying and assessing the environmental impacts of road projects. One of the most comprehensive of these has been developed by the World Bank (Tsunokawa and Hoban 1997). Many of these are in the form of checklists or matrices, and are quite detailed and extensive. Their main function is to serve as a kind of aide memoire for road planners to ensure that the kind of impacts described in the first part of this chapter are taken into account in the preparation of an EIA. Geographical information systems (GIS) also offer a powerful tool to assist with assessments, and provide a useful approach to visualizing the results of traffic and environmental models.

One of the inherent difficulties in assessing the impacts of road projects (as opposed, say, to energy or industrial projects) is that many of the impacts are outside of the ‘control’ of the road planner. These include effects such as noise and air quality. These are more directly associated with the quality, maintenance and use of motor vehicles, than of the physical infrastructure of the road itself. Despite this, there are a number of environmental models related to noise, emissions and air quality that can be integrated with transport demand models.

Despite the existence of various assessment methods and techniques, these formalized tools are rarely applied to actual EIAs. A United Nations task force on EIA noted that the most common ‘methods’ for identifying and assessing environmental impacts are ‘previous experience’ and ‘best professional judgement’.

EIAs can vary widely in terms of their length and level of detail. These depend upon the type of project and its complexity, as well as on the nature and sensitivity of the receiving environment. A successful scoping process will have ensured that

Guidelines and tools

Control of environmental impacts

Experience and judgement

EIA document preparation

Box 5.1 Typical components of an EIA

- Purpose and need of the project
- Project description
- Description of the existing environment
- Assessment of environmental impacts associated with construction, operation and maintenance
- Comparison of alternatives
- Description of mitigation measures and/or measures to enhance environmental benefits
- Outline of an environmental monitoring plan.

the EIA concentrates on the significant impacts, thus limiting its coverage to the most important issues. Very often EIAs, which have been prepared without scoping, turn out to be presented in voluminous and unwieldy documents. These tend to focus on description rather than assessment, and are less than helpful for decision-making. As a general rule, an EIA is made up of components such as shown in Box 5.1.

5.3.5 Valuation of environmental impacts

Once the impacts with a particular project have been identified (i.e. predicted) and their magnitude and significance assessed, the task remains of how to evaluate them. How should an environmental impact be measured? What is the cost of an environmental loss? What is the benefit of an environmental improvement?

Analysis

Cost–benefit analysis and multi-criteria analysis are the most commonly used methods of evaluating the impacts of infrastructure projects. The incorporation of environment in these tools, however, is still not commonplace. As Dom (1999) has pointed out:

The monetization of environmental impacts is possible to a certain degree, but there is no consensus as to the range of impacts that should adopt monetary values. The detailed methods and values used vary considerably between countries that adopt this approach. Problems mainly relate to the valuing of environmental impacts, as scientific knowledge is still incomplete and cost evaluation is often performed in a subjective or political manner.

Valuation

Ideally, environmental impacts should be evaluated in terms of legally binding, or generally accepted, environmental standards. However, with the exception of noise standards (and these vary widely from country to country), specific environmental standards related directly to road planning and construction rarely exist. Indeed, it is the very lack of such generally agreed environmental standards that has prompted the application of EIA itself as a planning tool. It is through an EIA that environmental impacts are actually evaluated. As that evaluation is most often a subjective one, the importance of access to information and public consultation has grown in importance.

5.3.6 Public consultation

Given the importance of transport in today's world, it is no surprise that the public has strong views about transport options, their costs and environmental impact. It is the subjectivity and political nature of the environmental effects of roads that make public consultation such an important part of any EIA process. Indeed, it has often been said that, despite engineering, economic or environmental constraints, roads most often follow the path of least political resistance.

Politicization

The involvement of the public in road planning is often difficult for a number of reasons. First of all, the definition of 'public' itself can be contentious. It normally involves a number of 'stakeholders' including road users, and representatives of industry, agriculture, households, business and services – often with conflicting interests. Consulting with these stakeholders on alternative road schemes often leads to polarization, with authorities siding with one side or another. Another problem with consultation is that often the stakeholders most directly affected by a project are not its main users. For example, farmers and rural communities suffer the negative impacts of severance, noise and air quality resulting from the construction of an urban by-pass, while city residents and businesses are the beneficiaries.

Stakeholders

As noted earlier, one of the best ways to ensure that all of the public's concerns are reflected adequately in the EIA is by scoping. This is important, not only to determine the specific range of issues to be considered in the assessment, but also to obtain agreement on those issues from the various parties who have a vested interest. Early agreement on the way in which the EIA should be carried out and on its content can help avoid delays at later stages in the decision-making process.

Scoping

5.3.7 Environmental controls and monitoring

Even the most comprehensive EIA (based on scoping, public consultation, and the most up-to-date methods for impact prediction and evaluation) is worth little, if the EIA report remains a 'stand-alone' document that is not integrated in overall project plan and its documentation. This is particularly the case with respect to mitigation measures. Although, in principle, an EIA could conclude that the negative environmental impacts of a proposed project outweigh any economic or transport benefit, this is rarely the case. An EIA report needs to be much more than the specification of mitigation measures related to the environmental impacts. A frequent shortcoming is that mitigation measures are not quantified in terms of their costs, nor is responsibility for their implementation assigned. When this is not done, successful implementation is questionable.

Quantification and responsibility for mitigation measures

Specific clauses need to be included in construction and supervision contracts to ensure that environmental mitigation measures are implemented correctly. Key issues that need to be addressed are those associated with site establishment and upkeep, construction work activities, and the restoration of the site after completion of the work. Each of these has particular impacts and mitigation options. Environmental clauses also need to be included in road maintenance contracts.

Contracts

As with mitigation, monitoring is an important element of the EIA process. Monitoring is important to ensure that the predicted environmental impacts actually occur and that, more importantly, the mitigation measures for dealing with them are

Monitoring

being implemented successfully. Monitoring data can also play an important role by refining the database and assumptions made in impact identification and assessment for use in future EIAs.

5.3.8 Strategic environmental assessment

Strategic decisions

While EIAs of individual road projects can often lead to environmentally improved design and the provision of environmental mitigation measures, they are, by definition, not able to examine the more strategic questions regarding the environmental effects of countrywide transport decisions. These include, for example, the environmental advantages and disadvantages of constructing motorways versus upgrading existing roads, and investing in other modes of transport such as rail.

Examples of SEA

The concept of strategic environmental assessment (SEA) has been developed as a way of addressing the environmental impacts of plans, programmes and policies. Examples of SEAs for different levels of transport planning can be found in several countries, particularly in Western Europe. The World Bank, which is involved in 'policy-based lending' and finances large structural adjustment schemes, have carried out a number of SEAs (usually called 'regional environmental assessments') covering geographical regions rather than investment sectors such as transport.

Transport corridor planning

On the local and regional level, transport SEA is increasingly integrated with, and performed as a part of, the land-use planning process. The other main focus of transport SEA has been on transport corridor assessment. The European Commission, for example, is undertaking a pilot SEA of the whole multi-modal Trans European Network and for various transport corridor assessments in co-operation with member states. The recent adoption of the European Union *Directive on Strategic Environmental Assessment* is also likely to increase attention on carrying out SEAs in the transport sector in Europe.

5.4 Concluding remarks

Current situation

The environmental impacts associated with road projects are well known and well documented. Procedures and regulations for carrying out EIAs are in place in a large number of developing countries and most emerging countries, and this has been encouraged by multilateral financial institutions and bilateral aid agencies. There is abundant guidance available for carrying out EIAs in a thorough fashion. However, the practice of EIA is often more of an art than a science. This is due to the uncertainty of the predictions involved, and to the political nature of decisions affecting the human and natural environment. It is for these reasons that scoping and follow-up public consultation are so important in any EIA process.

Assessment teams

A key element for carrying out a successful EIA is the personnel involved in the assessment team. The specific make-up of any team, in terms of the number and type of particular disciplines represented, can vary greatly from case to case depending on the particular project, its characteristics and type of project environment. Regardless of the specific configuration, however, it is important that EIAs carried out in developing and emerging countries are undertaken by teams made up of both international and local experts. Too often, Western consultant firms, with little or no knowledge of the local environment, have carried out assessments, which have

proven to be little more than translations of local studies. EIAs should be based on up-to-date environmental information, which can best be provided through local experts, together with the best available methodologies and approaches developed internationally. In such circumstances, EIAs can make a positive contribution to environmentally sustainable transport.

References

- Dom, A. (1999) 'Environmental impact assessment of road and rail infrastructure', in Petts, J. (ed.), *Handbook of Environmental Impact Assessment, Volume 2 Environmental Impact Assessment in Practice: Impact and Limitations*, Oxford: Blackwell Science, 331–350.
- EBRD (2003) *Environmental Procedures*, London: European Bank for Reconstruction and Development.
- Tsunokawa, K. and Hoban, C. (1997) 'Roads and the environment: a handbook', *World Bank Technical Paper No. 376*, Washington, DC: The World Bank.

Planning methods

Knud Rask Overgaard

6.1 The nature of planning

Planning and programming

A plan is a proposal for undertaking future activities. Planning can be a detailed process relating to activities to be undertaken in the relatively near term, and this is sometimes known as ‘tactical planning’. More normally, within the road management context, the term ‘programming’ is used for this (see Chapter 19). By contrast, ‘strategic planning’ involves an analysis of the road system as a whole, typically requiring the preparation of long-term estimates of expenditures for road development and conservation under various budgetary and economic scenarios. Planning is often undertaken by a planning or economics unit.

Planning and policy

The results of the planning exercise are usually of most interest to senior policy makers in the road sub-sector, both political and professional. There is a clear link between planning and policy (see Chapter 2). A plan provides the strategy for putting a policy in place – it provides the dimensions of time and responsibilities for action, whereas policy is essentially a ‘statement of intent’. Conversely, policy can also be the result of a planning exercise. Planning enables different scenarios to be examined to see which are the most favourable or desirable for inclusion in the policy.

Development aims

The development of the road system should be planned so that it will support the general development aims of the country, including such things as economic growth, uniform regional development and full employment (Replogle 1991). Based on these general aims, more specific aims for the transport sector may be developed. At the strategic level, the network focus means that road planning often needs to be undertaken in conjunction with transport sector planning. The broad aims of a society are normally fairly clear to the government and politicians, and understood in general by the public. The aims of a transport policy, however, are often not very well defined unless a published policy document has been produced. Even then, the policy may only be defined through general phrases such as ‘inexpensive’, ‘fast’, ‘reliable’ and ‘safe’ transport.

Issues

Some of the most important issues to be determined by transport planning and road network development are

- How much should be invested in public transport infrastructure, and to what extent should the level of investment be determined by the economic rate of return and by social/regional considerations?
- How should public transport investments be divided between roads, railways, ports, airports, etc.?

- Should emphasis be on primary, secondary or tertiary roads?
- What kind of road projects should be given priority; for example, those increasing capacity, those reducing costs, those increasing the level of service, those improving safety or those improving the environment?
- How should financial allocations be divided between new construction, rehabilitation and maintenance of roads?

The answers to these questions are often elaborated in the form of a transport plan or road development plan. These are likely to consist of a long-term strategic plan, a medium-term development programme and an annual works programme, which is updated every year. The national road plan represents a framework for more detailed plans and studies, including feasibility studies, design and implementation of specific road projects, and for the co-ordinated implementation of other activities contained in the plan. Planning needs to be a continuous process, with on-going recording of developments, so that plans can be updated as conditions change.

Published
plans

6.2 Strategic planning

Strategic planning is used to assist the development of policy. Strategic planning might involve the determination of appropriate standards that should be used on the different hierarchies of roads within the network. It could also involve an investigation of the impact on future road conditions of applying different budget levels. Thus, typical outputs from the strategic planning process might include

Policy
development

- Projected annual capital and recurrent budget requirements to meet road administration standards over a user-defined period into the future (see Figure 6.1).
- Projected road conditions resulting from the application of pre-defined annual budgets for a user-defined period into the future (see Figure 6.2).
- Projected road administration costs and road-user costs for pre-defined standards, or annual budgets, for a user-defined period.
- Incremental net present value (see Chapter 7) of adopting one set of standards compared with another, or of adopting one particular stream of annual budgets compared with another.

Strategic planning involves an analysis of the entire road network, typically with road sections grouped by traffic level and road condition. Individual sections are not identified, and the network is characterized, for example, by lengths of road, or percentages of the network, in various categories defined by parameters such as road class or hierarchy, traffic flow or congestion or pavement type. Alternatively, the network may be characterized by the proportion of the network in a particular condition state. Strategic planning requires life cycle analysis of the cost of construction, maintenance and the road user across the network. The steps involved in the analysis are illustrated in Figure 6.3. This shows that each step requires that basic data exists about the road network. The Highway Development and Management Model, Version 4 (HDM-4) is a useful tool for carrying out this analysis (see Chapter 21).

Analysis

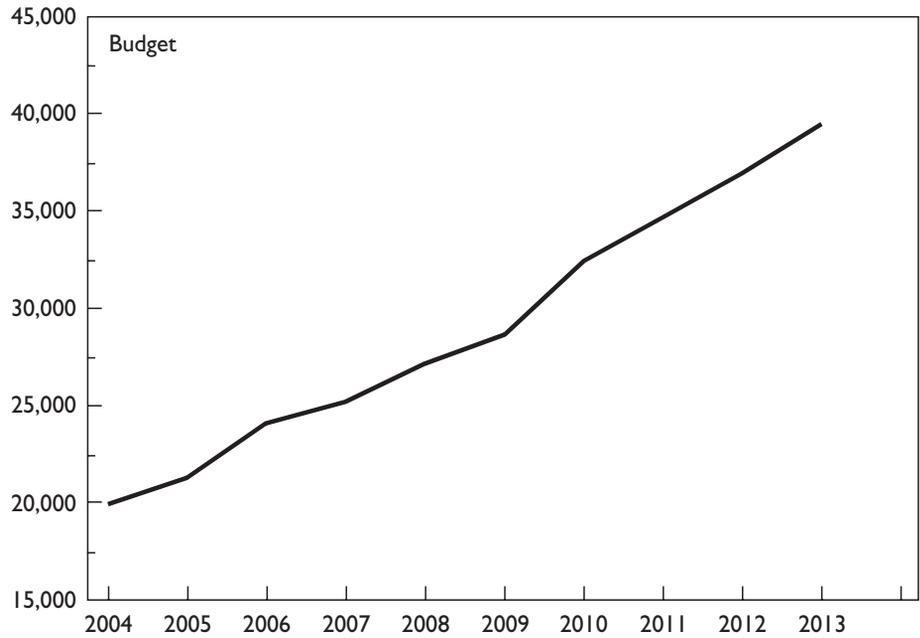


Figure 6.1 Example of budget required to maintain a given road condition.
Adapted from: TRL (1998).

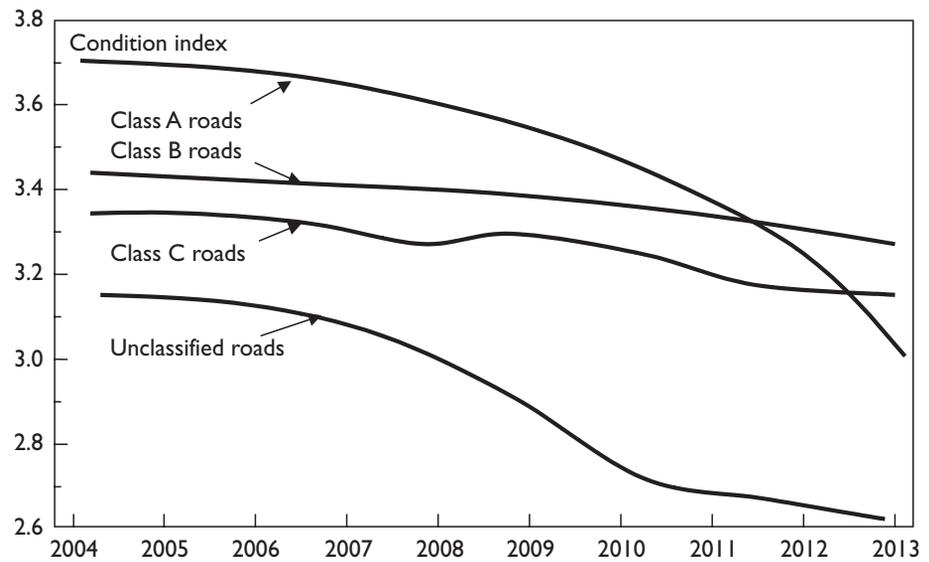


Figure 6.2 Example of projected road conditions resulting from the application of a fixed budget.
Adapted from: TRL (1998).

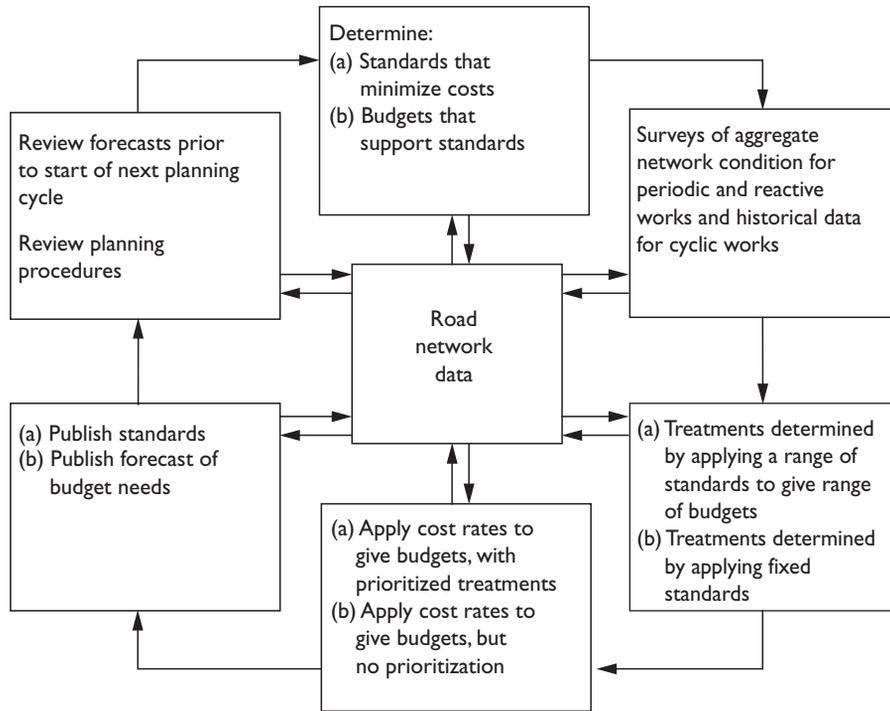


Figure 6.3 Steps in strategic planning for two examples of decision.

Adapted from: TRL (1998).

6.3 Physical network planning

Physical network planning is undertaken to identify where long-term road development should be targeted. The starting point for developing these plans is the existing road network. An inventory of this is needed in terms of its extent, standard, condition, utilization, etc. Other important inputs are the relevant objectives and policies established by government, the expected future transport demands, and the available resources – primarily finances.

There is no definitive method for determining road development requirements. The planning methodology applied in a particular country should be adapted to the actual conditions and problems at the time of planning (Thomson 1983). An iterative procedure is usually adopted, by which an initial network proposal is gradually improved by examining the consequences of marginal network changes. This process may be preceded by a ‘scenario analysis’ that compares a small number of widely different network proposals. This is particularly useful when strong intermodal competition exists, in which case the approach may be used to establish the approximate demand for road transport. The road development plan can then be worked up in more detail.

The data needed for quantitative planning of a road system is often not readily available in developing and emerging countries. Thus, the first approach to national road planning on an analytical basis may have to be based on rough and partly estimated data.

Information needs

Methodology

Data needs

Estimating future traffic is still an imprecise – but essential – art. Since transport investments commonly have long lives, decisions to make such investments inherently involve long-term forecasts. It is clearly preferable to be explicit about the underlying assumptions than to leave them unstated.

(Adler 1987)

An analytical approach often highlights problems that would otherwise be neglected, and it gives rise to a usually much needed systematic collection of data concerning roads and traffic. Geographical information systems (GIS) can facilitate the planning process (Denno 1993), but do not substitute for a robust methodology.

Establishing the network

Due to the uncertainties associated with traffic forecasting (see Chapter 3), and the long-term nature of road investments, the road development plan should not cover more than a 10–15 years period. The network likely to be in place at the end of this period may be established by adding the following four types of projects to the existing network:

- Current and committed projects, including those for which a construction contract has been signed.
- Projects that are natural extensions of existing and committed projects to complete a route or complete the basic structure of the road network.
- Projects considered important from a development or strategic point of view.
- Other technically feasible projects, which have potential economic value, but are not essential to the structure of the road network.

Candidate roads

A useful step in identifying candidate roads for the plan is to establish a map indicating major transport desire lines between the future main activity areas (see Figure 6.4). By comparing these corridors of potentially high transport demand with the existing network configuration, proposals for new road links may emerge (see Figure 6.5). By assigning the estimated future transport demand to the existing road network, it may be possible to identify corridors that will require capacity increases and links that will require upgrading in the future.

Road standards

The road standards to be applied to the potential new networks should take account of the expected future traffic volumes, since minimization of total transport costs is likely to be an important objective of the road development plan. However, the opportunity should be taken during the planning process to investigate geometric, pavement design and maintenance standards. For all of these, an optimum standard will exist that minimizes total transport costs, as illustrated in Figure 6.6. Often government standards for roads are too high for given traffic levels and resources. For example, for rural access roads, simple improvements to a dirt track are often sufficient to meet transport needs at a low cost, rather than the provision of a higher-standard road. Determining the optimum standard is a strategic planning exercise.

Incremental analysis

The network of existing roads and current and committed projects provides the basis for an incremental analysis in which candidate projects are added gradually to the basic network. The feasibility of an incremental expansion of the network is tested by comparing road-user savings with the construction and maintenance costs required to implement the expansion. In this way, the optimal plan may be determined, and important indications received about appropriate priorities for network development.

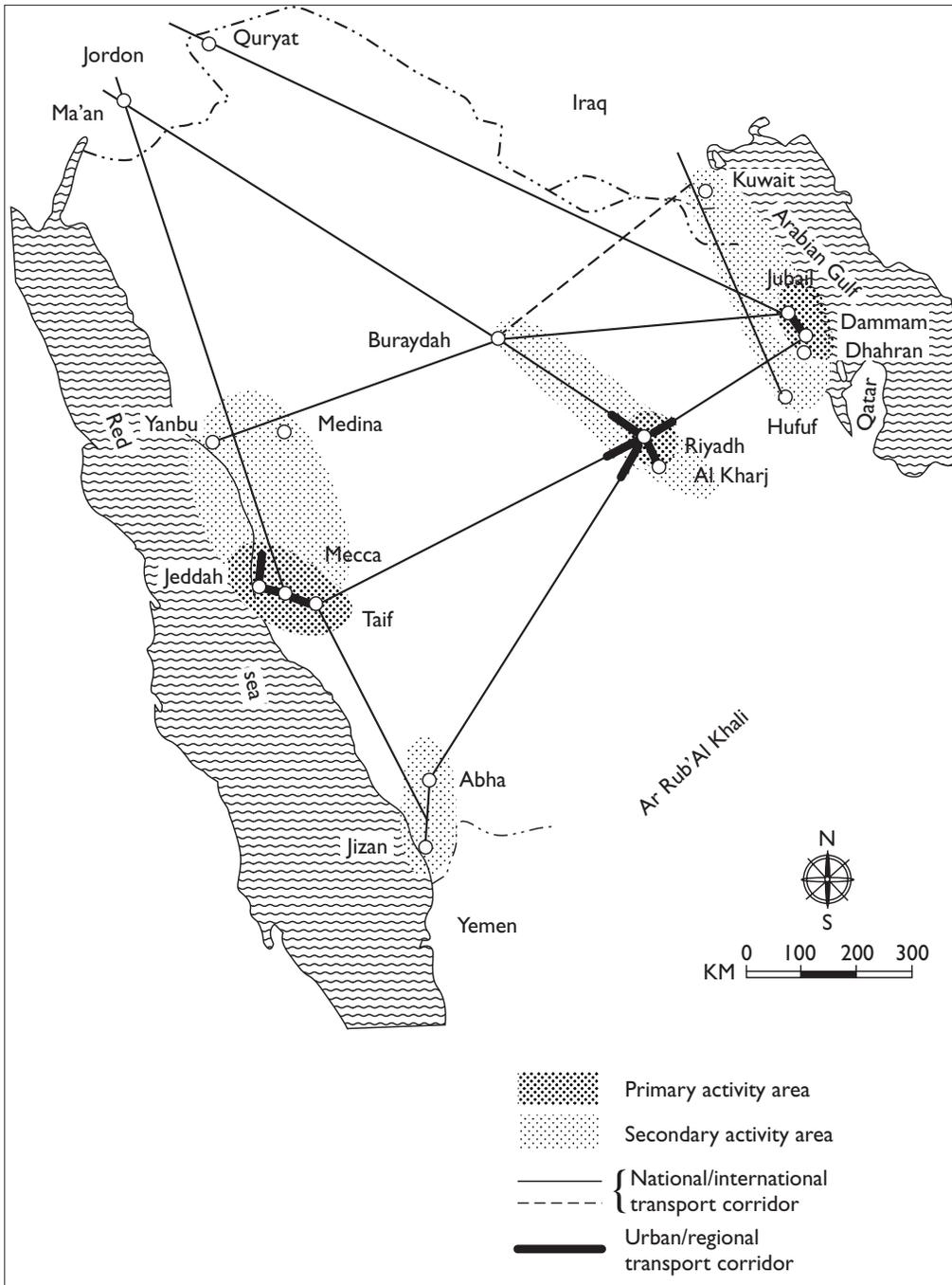


Figure 6.4 Example of main activity areas and major transport desire lines.

The analysis discussed here has focused on the economic aspects of primary and secondary roads. In addition, planning is needed for rural access roads, which provide access to adjacent land for motorized and non-motorized vehicles. They have a significant social role. The planning of rural access roads is discussed later.

Rural access roads

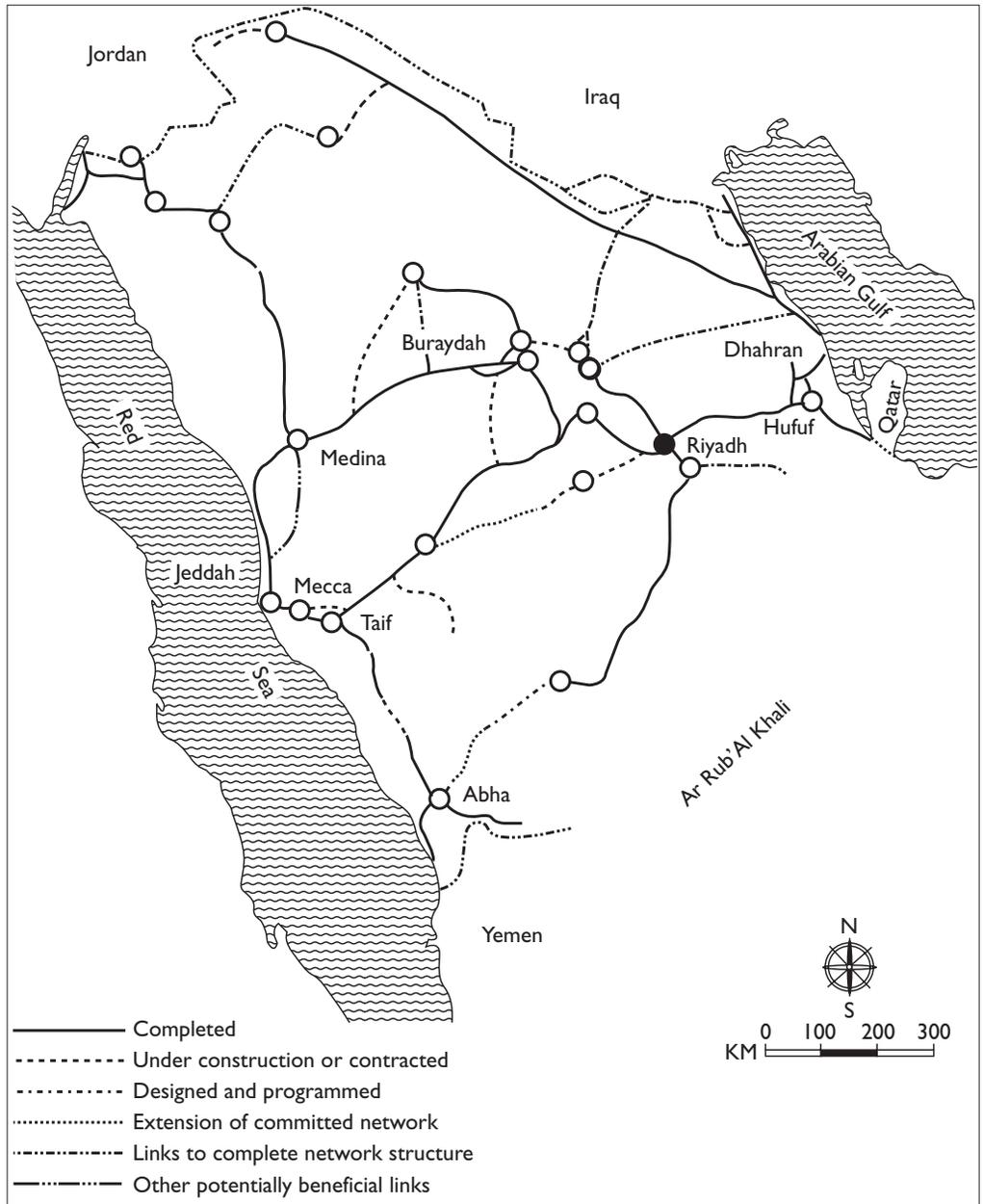


Figure 6.5 Example of existing and potential new road links.

6.4 Transport demand forecasting

6.4.1 Traffic forecasting for individual roads

Need for
traffic
forecasts

Information on current traffic and estimates of future traffic are necessary to prepare optimal designs for improved or new roads and to make efficient plans for development of road networks. The existing and the diverted traffic together are considered the current traffic. The future traffic further comprises normal traffic growth, generated traffic and development traffic (see Figure 6.7).

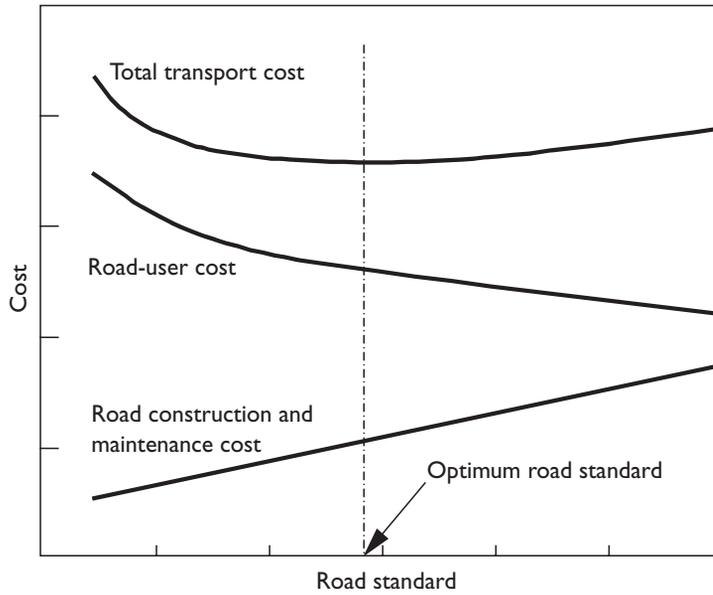


Figure 6.6 Optimum road standard.

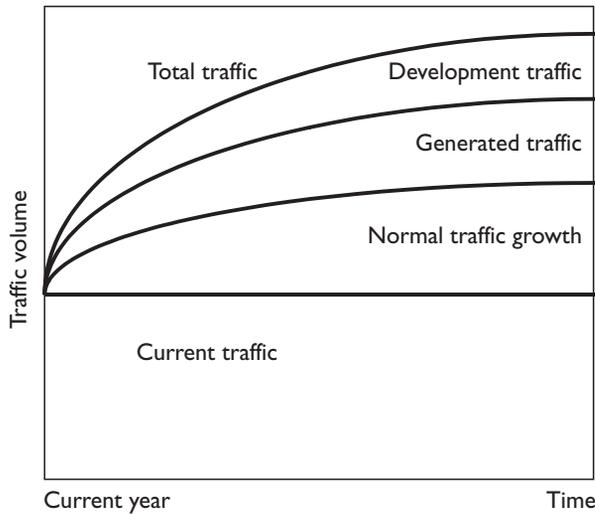


Figure 6.7 Traffic development on an improved road.

When the influence of a road improvement or a new road is confined to a narrow road corridor, it is sufficient to base traffic forecasts on the existing road traffic (in the case of an improvement), and the traffic which may be diverted from the immediate parallel roads upon opening of the improved or new road.

Corridor forecasts

If the future economic development is expected to continue as in the past, then the normal traffic growth may be estimated by extrapolation of time series data of traffic volumes on the existing and/or parallel roads. The following growth curves are commonly used.

Growth curves

Linear growth: $y = a + b \cdot t$

Exponential growth: $y = c \cdot e^{dt}$

Logistic growth: $y = \frac{Y_s}{1 + f \cdot e^{-gt}}$

where y is the traffic volume; t the time variable (e.g. number of years after base year); e the base of natural logarithms (i.e. 2.7183); Y_s the saturation level (asymptote); and a, b, c, d, f, g are parameters, which may be developed from time series data (or otherwise established).

Traffic and GNI

As an alternative to extrapolation, traffic growth can be related linearly to gross national income (GNI) (see Chapter 1). This is preferable, since it takes into account, explicitly, changes in overall economic activity. However, using this relationship for forecasting requires a realistic forecast of GNI, which is not always available. If it is expected that a particular component of the traffic will grow at a different rate than the rest, then this should be identified and dealt with separately.

Generated traffic

‘Generated traffic’ arises because a journey becomes more attractive as a result of the new or improved road. It is only likely to be significant in those cases where the road investment brings about large reductions in transport costs. In the case of a minor improvement within an already developed road system, generated traffic will be small and can normally be ignored. Similarly, for projects involving the improvement of short lengths of rural roads and tracks, there will usually be little generated traffic. However, in the case of a new road allowing access to a hitherto undeveloped area, there could be large reductions in transport costs as a result of changing mode from head-loading to motor vehicle transport and, in this case, generated traffic could be the main component of future traffic flow.

Development traffic

‘Development traffic’ consists of the additional trips to areas where the development has been speeded up due to a new or improved road. This traffic, which usually develops over a long period of time, depends upon the extent to which authorities allow development to take place, and the actual land use alongside the road. In practice, it is difficult to distinguish between generated traffic and development traffic, so the two are usually considered together as ‘generated traffic’.

Price elasticity

Demand relationships can be used for forecasting generated traffic. The price elasticity of demand for transport measures the responsiveness of traffic to a change in transport costs following a road investment. As costs reduce, demand increases, and vice versa. The price elasticity is defined as the ratio of the percentage change in travel demand to the percentage change in travel cost. Evidence from studies in developing and emerging countries suggest a range of between -0.6 and -2.0 for the elasticity, with an average of about -1.0 . This means that a 1 per cent decrease in transport costs leads to a 1 per cent increase in traffic (TRRL 1988). The available evidence suggests that the elasticity of demand for passenger travel is usually slightly greater than this average of about -1 . However, the elasticity of demand for goods is much lower (-0.1 to -0.5), and depends on the proportion of transport costs in the commodity price. Demand elasticity is discussed further in the Chapter 7.

Door-to-door travel costs

All calculations should be based on door-to-door travel costs, and not just on that part of the costs incurred on the road under study. Generally, this implies that the reduction in travel costs and increase in traffic will be smaller than measurements on the road alone suggest.

6.4.2 Traffic forecasting for road networks

In order to plan the development of the road network in an optimal way, it is necessary to assess future transport demand. An estimate is needed of the transport demand between the various parts of the area under study to predict the possible future traffic volumes. For this purpose the area is divided into zones, which generate or attract traffic. Table 6.1 shows how the transport demands of an area, divided into n zones, are expressed in the form of a traffic distribution matrix, where T_{i-j} is the number of trips generated by zone i and attracted by zone j ; G_i the total number of trips generated by zone i ; and A_j the total number of trips attracted by zone j .

Traffic matrix

Different matrices may be used to express future transport demand in terms of different aspects, such as total passenger trips by all modes, car trips for all or specific travel purposes, total goods volumes by all modes, and truck trips with selected commodities.

The existing trip distribution needs to be established, in terms of an origin–destination matrix, by means of traffic surveys or application of a model with subsequent calibration against traffic counts (Cascetta and Nguyen 1988). Future trip matrices can be developed by using a growth factor model. The growth in traffic generated/attracted by each zone is estimated on the basis of an analysis of the expected development of traffic generating sources. For each zone, separate growth factors may be developed for generated and attracted traffic. However, if the O–D matrix for existing traffic does not distinguish between traffic generated and traffic attracted by a zone, a composite growth factor (F) will have to be established for each zone. In this case, the average growth factor model estimates the future traffic (T_{ij}) between two zones i and j starting from the formula:

Growth factor model

$$T_{ij} = t_{ij} \frac{F_i + F_j}{2}$$

where t_{ij} is the observed traffic; and F_i, F_j the growth factors for total traffic from/to the zones i and j , respectively.

Table 6.1 Matrix for trips between n zones

From zone	To zone						$\sum_j T_{i-j}$	
	1	.	.	j	.	.		n
1	T_{1-1}	.	.	T_{1-j}	.	.	T_{1-n}	G_1
.
.
i	T_{i-1}	.	.	T_{i-j}	.	.	T_{i-n}	G_i
.
.
n	T_{n-1}	.	.	T_{n-j}	.	.	T_{n-n}	G_n
$\sum_j T_{i-j}$	A_1	.	.	A_j	.	.	A_n	$\sum_j T_{i-j}$

The calculations usually require an iterative procedure to balance the system (Overgaard 1966). The average growth factor model does not take into account explicitly future changes in access between the various zones. Therefore, it is more applicable to situations with limited or uniform transport improvements than to major 'spot' improvements of the road network.

Growth factor The passenger traffic growth factor of a zone may be calculated as follows:

$$F = \left[\left(\frac{I \cdot E}{100} + 1 \right) CP \right]^n$$

where F is the growth factor for total passenger traffic to/from the zone; I the annual growth rate (per cent) for per capita income in constant prices; E the passenger transport demand–income elasticity; CP the compound population growth rate per year (e.g. 1.03 if population grows by 3 per cent per year); and n is the number of years between survey year and forecast year.

Applicability The growth factor model is applicable mainly to short- and medium-term forecasts. Therefore, variations in I , E and CP over time should be taken into account for forecast periods of more than, say, 10 years (i.e. $n > 10$) by calculating F in two or more steps, if a growth factor model is to be employed. For medium- and long-term forecasting, the 'gravity model' presented in Sub-section 6.4.3 may be a more relevant tool, because it has greater ability to reflect changes in geographic development patterns and transport accessibility. If modal competition is significant in the study area, the modal split may have to be taken into consideration directly when estimating future road traffic volumes between zones. Finally, these volumes should be assigned to road sections in the network to estimate future traffic flows.

Four-steps model The above description indicates a sequential traffic forecasting model consisting of the following four steps, or sub-models:

- 1 Traffic generation and attraction by zone
- 2 Traffic distribution (matrix estimation)
- 3 Modal split (when required)
- 4 Traffic assignment.

This model is relatively simple, and has widespread applicability. It is actually two models: one for passenger transport and one for goods transport. Other models combine some of the four steps (Overgaard 1966).

6.4.3 Passenger transport forecasting

Model development The estimation of the future demand for passenger transport is usually based on an analysis of existing travel behaviour. This results in models that explain passenger trip-making in relation to socio-economic variables, such as demographic factors, personal incomes, transport demand–income elasticities, etc. If it is not possible to collect local data on the likelihood of trips, then it may be necessary to use rates from other sources that are applicable to the situation in question.

Traffic estimation Initially, it is necessary to decide whether road traffic may be estimated independently, or whether competition from other modes needs to be taken into

account explicitly. When road traffic can be forecast directly, the estimate is often determined separately for cars or high-occupancy vehicles, such as buses and para-transit vehicles. The number of passengers is converted to equivalent vehicles using average occupancy rates. As trips for different purposes exhibit different distribution patterns and growth rates, it is necessary to decide whether all trips can be grouped, or whether they should be divided by purpose, such as business trips and leisure trips.

The number of passenger trips generated/attracted by a zone may be estimated as the population of the zone times a trip rate. A more advanced method distinguishes between urban and rural populations, and applies different trip rates to the two groups. The trip rates usually also vary with real income per capita. This should be taken into account when the future traffic generation/attraction by zone is calculated. This is done by applying appropriate transport demand–income elasticities and economic growth assumptions to adjust the observed trip rates. Elasticities between 1 and 2 are common depending, among other things, on the degree of development.

The number of passenger trips between the zones may be calculated using a constrained gravity model:

$$T_{ij} = a_i \cdot G_i \cdot b_j \cdot A_j \cdot f(d_{ij})$$

where T_{ij} is the number of trips generated by zone i and attracted by zone j ; G_i/A_j the total trip generation/attraction by zone i/j ; $f(d_{ij})$ a function which expresses the resistance that has to be overcome to travel between the zones i and j , for example, $f(d_{ij}) = d_{ij}^\alpha$, where d_{ij} is the travel time between the zones i and j ; α a parameter estimated on basis of an analysis of the present transport supply and the observed trip distribution; and a_i, b_j calibration factors to ensure that the row and column sums in the trip distribution matrix correspond to zonal trip generations and trip attractions, respectively. The future trip distribution is calculated by iteration (Overgaard 1966).

If the trip matrix includes trips by more than one transport mode, a ‘modal split’ calculation needs to be undertaken for zone pairs. When two modes only are competing, the use of modal split percentages relating to transport cost or time ratios may be sufficient, as shown in Figure 6.8.

If several public transport modes are available between two zones, traffic may be divided between these on basis of the following expression:

$$T_m = k \frac{f_m}{c_m t_m}$$

where T_m is the traffic by mode m ; f_m the frequency of mode m ; c_m the transport fare for mode m ; t_m the travel time for mode m ; and k a parameter which ensures that traffic by all modes corresponds to total traffic between the two zones.

The logit model may be used to determine mode and route distribution simultaneously. The model is usually based on individual travel behaviour and not on average figures for zones. However, the development of the model is somewhat complicated and requires data from passenger interviews and detailed data on relevant transport alternatives (Ben-Akiva and Lerman 1991).

Trip
generation

Traffic
distribution

Modal split

Public
transport

Logit model

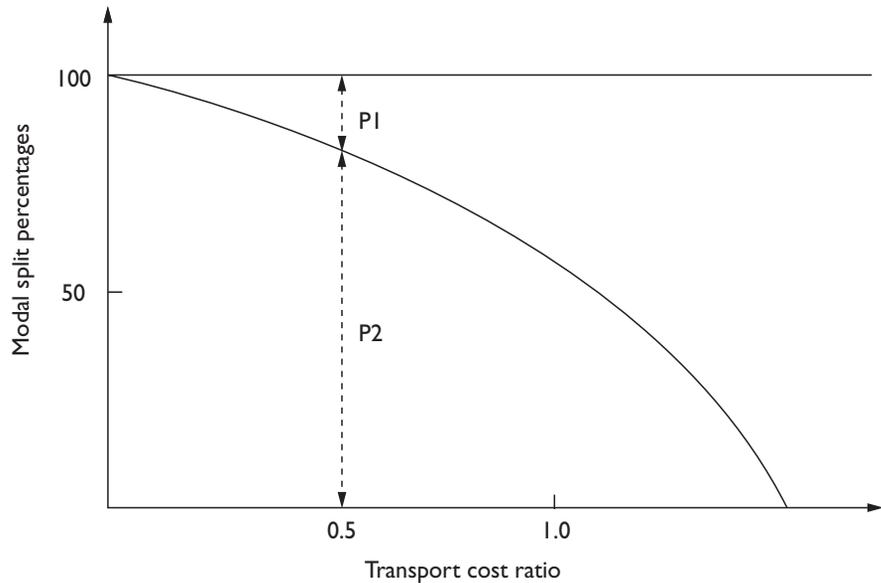


Figure 6.8 Modal split model.

Traffic assignment

The traffic matrix may be assigned to the road network alternatives using the ‘all-or-nothing’ principle, which allocates all traffic between two zones to the optimum route, or by multiple route assignment (Overgaard 1966). The optimum route may be the fastest or the cheapest route, or be based on a combination of travel time t and distance d , that is

$$a \cdot t + b \cdot d$$

where a and b are parameters estimated from observed route choices.

Travel time components

The travel time used in the analysis should include all components of the door-to-door trip. Thus, for a journey by public transport, travel time may be calculated as follows:

$$t = a \cdot t_1 + b \cdot t_2 + t_3 + c \cdot t_4 + t_5 + d \cdot t_6$$

where t is the door-to-door weighted travel time; t_1, t_6 the time walked to/from public transport; t_2 the time spent waiting; t_3, t_5 the travel time in public transport vehicles; t_4 the transfer time between routes; and a, b, c, d parameters which express the inconvenience the travellers experience in connection with the various time components; normally the parameters are higher than unity, and may be estimated from observed route (or mode) choices.

Example

Figure 6.9 presents a possible methodology for passenger traffic forecasting, which applies a growth factor model. The methodology may include or exclude a modal split calculation, as appropriate. It may also be applied directly to car or passenger vehicle traffic. Traffic assignment is accomplished by means of the all-or-nothing principle using fastest (shortest) routes. Local road traffic must be added to

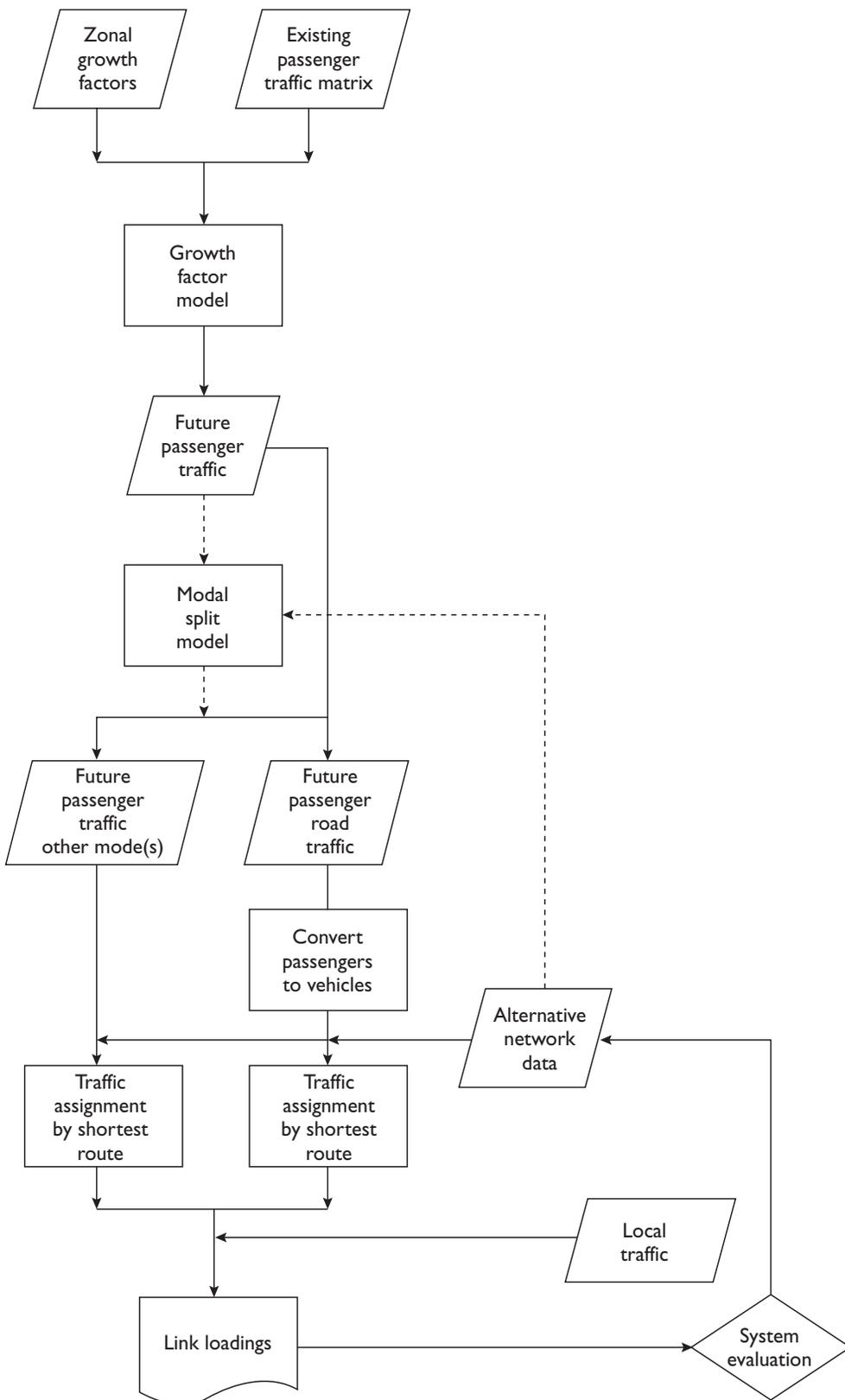


Figure 6.9 Methodology for passenger traffic forecasting and system evaluation.

the interzonal flows assigned to each link, and taken into account when alternative road networks are evaluated.

6.4.4 Goods transport forecasting

Commodity classes

Forecasts are undertaken separately for different groups of commodities being transported, such as agricultural products, petroleum products, building and construction materials, industrial products, consumer products. This is to reflect the different transport characteristics, growth and distribution patterns. Usually fewer than 10 groups will suffice, and sometimes only total truck traffic needs to be considered.

Traffic generation

The traffic generation/attraction of each zone is estimated on the basis of the expected future 'surplus/deficit' of each commodity. 'Surplus' is the production (plus import) less local consumption within the zone, and 'deficit' is the consumption (plus export) less the amount supplied locally. Thus, the surplus/deficit figures indicate the tonnages to be transported between the zones. Intrazonal traffic has to be determined separately. The estimation is accomplished by analysis of the regional economy, including historical growth rates and future growth potential. This needs to take into account land usage, plans for agricultural, industrial and construction projects, employment, population and personal income development, propensities to consume, etc.

Traffic distribution

Models similar to those for passenger traffic may also be used to estimate future goods traffic distribution, modal split and route assignment. Alternatively, a 'transportation model' may be used.

Transportation model

Transportation models reflect intermodal competition by matching, for each commodity, the surplus/deficit of the zones in such a way that total transport costs are minimized. The models require that the various transport modes are combined into one network, with the possibility of interchange between modes. They consider the minimum cost mode and route between each pair of zones. The least cost mode/routes are found by applying a 'shortest' route algorithm to the multi-modal transport system, using linear programming (Dantzig 1951). Assignment may be completed, as before, by means of the all-or-nothing principle, or using multiple routes.

6.5 Plan development and implementation

Plan and programme

A long-term development plan has to be turned into a works programme. This needs to take into account the limited resources, and particularly the finances, available to the road sector in most developing and emerging countries.

Equity

Improvements to roads of different classes affect different population groups in different ways. This adds further complications to the problem of defining the best possible road development programme within given budget constraints. It may be necessary to strike a balance between overall objectives, such as economic growth and efficiency, and regional and individual equity. This is illustrated in Figure 6.10. Curve 1 shows, for different traffic volumes (or road classes), the level of service (or the road standard) that would be the most economic for society if no budget constraint existed. Curve 2 represents a situation where a minimum level of service has been fixed for roads with low traffic volumes, but the available budget is insufficient

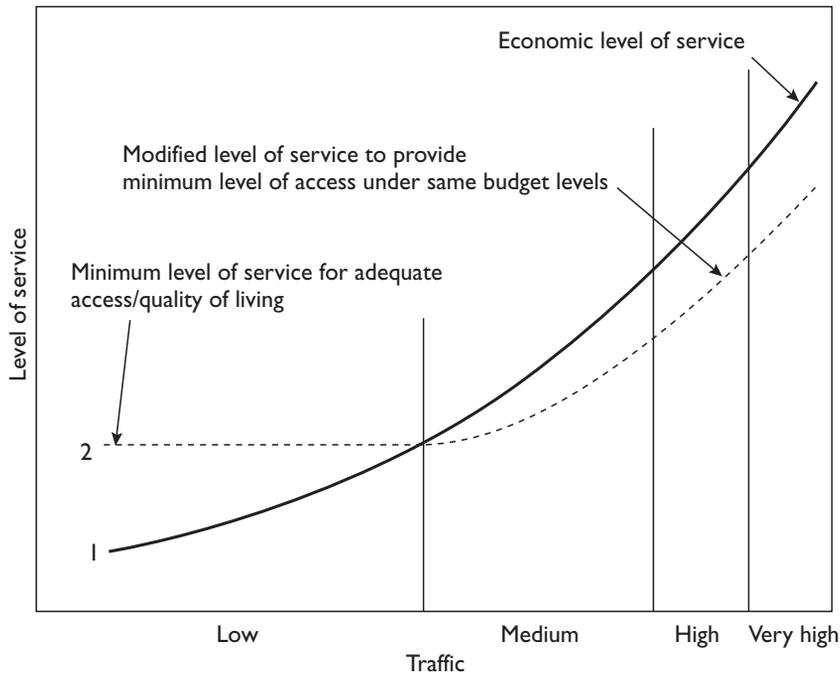


Figure 6.10 Effect of equity objective on economic efficiency standard.

to provide the high level of service that would be economical for roads with high traffic volumes. Thus, there will usually be a political dimension to planning decisions in addition to economic development and budgetary criteria.

It is also important to avoid overinvesting in particular road classes, or in specific geographical regions, by adopting road standards or road network densities before they are justified by traffic levels. This may deprive other road links or areas of needed improvements.

The projects with highest returns, within each road class, are selected from the long-term road plan for early implementation. Project returns can often be improved by implementing proposed developments in stages, matching better the demand over time. In this respect, the proper timing for paving roads is of particular importance in many countries. The optimal timing of implementation is also relevant for asset preservation and maintenance projects.

A purpose of planning is to identify the budget needed to implement all economically viable projects. This enables the real needs of the road sector to be demonstrated, and identifies the returns that society could expect from investments compared to those in other sectors. However, there will always be a limit to the available budget for the road sector, so it is important to establish the funding needed and likely to be available over a period of, say, 3–5 years. Practical considerations also need to be taken into account concerning implementation capacity. This is to ensure that the programme prepared will be realistic, and can serve as a proper basis for future implementation, and not just as a ‘wish-list’.

Figure 6.11 shows an example of the investment programme required to implement a 15-years roads development plan, although maintenance and rehabilitation

Over-investment

Stage construction

Budget requirements

Investment programme

costs are not shown. The programme is divided into three 5-year periods, with the first one being further subdivided into 1-year periods. The current year is included for comparison. The budget is divided into separate headings for primary/secondary and tertiary roads. The budget has one component to cover investments in projects identified by a specific planning study, and another to allow for additional projects that might be identified later. It should be noted that only a small number of primary and secondary roads is expected to be added to the plan, whereas a larger number of the tertiary roads are still unidentified. Furthermore, this share of tertiary roads increases towards the end of the planning period.

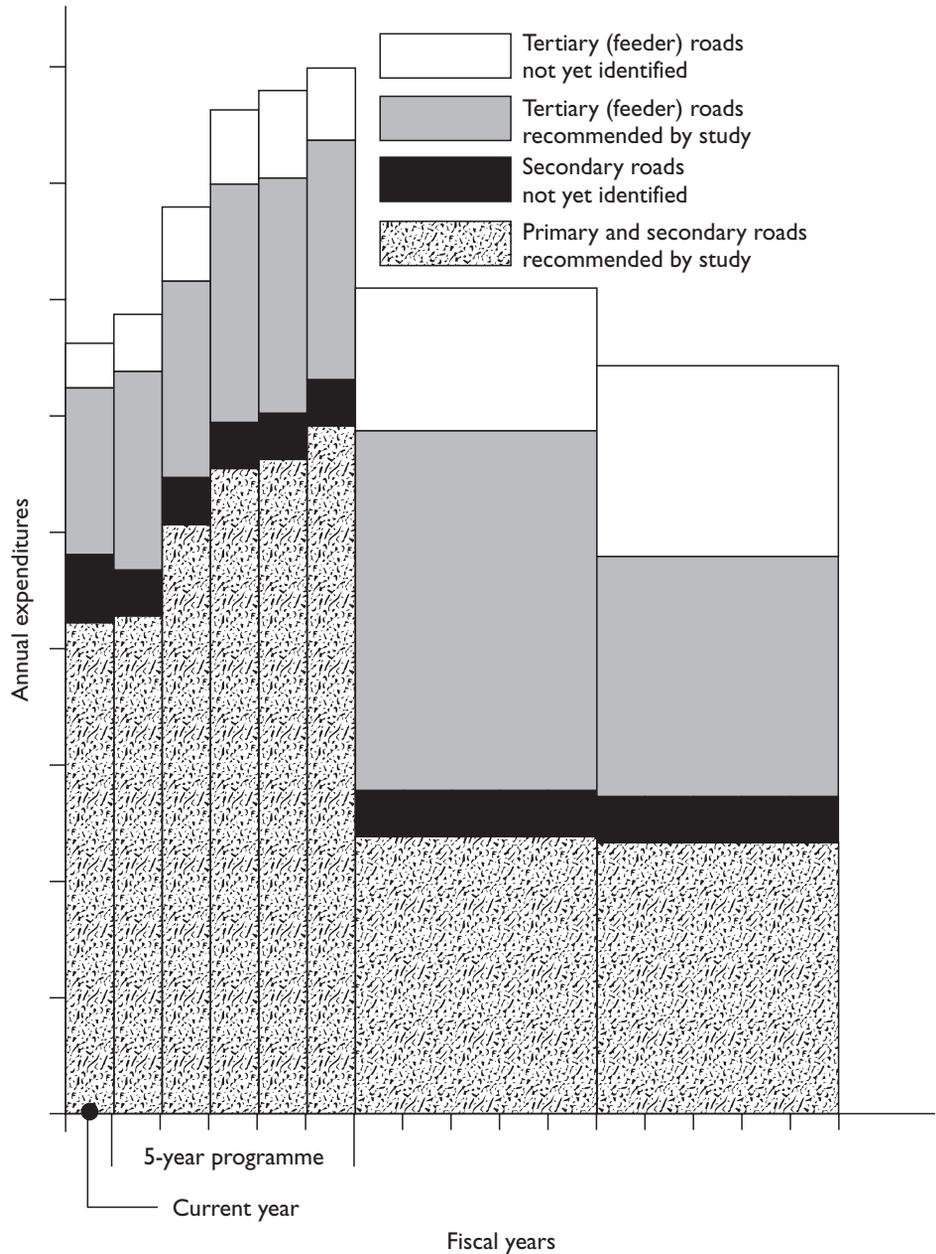


Figure 6.11 Roads investment programme.

Effective planning, programming and budgeting require access to information of an appropriate level of detail. Most countries now use computerized road management systems to assist with these processes. Some general requirements for systems are

Road management system

- To ensure that investments in the inter-urban road network are allocated to programmes, regions and links to produce the maximum economic growth throughout the nation; this can be achieved through the application of sound economic principles to the investment process.
- To balance the needs of national objectives of regional and individual equity, recognizing the tension that exists because of the requirement for maximum economic growth and the establishment of minimum serviceability levels.

Other system requirements are

- The ability to deal with frequent major fluctuations in budgets and policies, which necessitate major programme changes to be carried out in short time scales; the system should be capable of producing revised plans and programmes at short notice – this is easier with systems that are both integrated and computerized.
- Plans and programmes must be justifiable economically, so that financing from international lending agencies can be sought; it is helpful if the analysis of the investment programmes is based on the use of a model such as HDM-4; however, the model must be suitably adapted and calibrated for local conditions.
- Systems should be easy to operate to facilitate their use in provincial offices.
- The system should only be used as a tool to assist in developing plans and setting programmes, as there are often national objectives or provincial priorities that will take precedence over the programme determined by the system; flexibility to accommodate these changes in the process is necessary.

The use of computer models and systems to assist with the planning process makes it easy to analyse scenarios. Studies can be made of the alternative sets of allocation distributions, such as asset preservation, network development, training, studies, special projects, calamity, etc. The results of the scenario analysis, under various budget and regional block allocation constraints, is the starting point for the multi-year programming process and annual budgeting.

Scenario evaluation

The multi-year programming process takes programmes and projects from the long-term plan scenario, and distributes them over time. A typical approach is to place potential projects into blocks of five years each, in accordance with their economic ranking. The next step distributes projects in the first 5-year block by year, considering issues such as development potential, route integrity, social development, environmental considerations, etc. The level of preparedness of a project, in terms of feasibility study and design, may also influence this. The multi-year programme is used to produce the final annual budget. Thus, the programming activity follows a stepwise flow, as illustrated by the following:

Multi-year programming

- 1 Produce a preliminary multi-year programme and use this to prepare a list of projects for incorporation into the next budget.

- 2 Finalize the programme for the next budget year, and update the multi-year programme for the subsequent years.
- 3 Update the programme with ongoing/committed projects, in terms of current expenditures and status of completion.
- 4 Update cost figures for other ongoing/committed projects in the forward programme.
- 5 Update the integrated long-term plan with the new data and information.
- 6 Generate next year's preliminary multi-year programme and produce a preliminary list of next year's projects.

Monitoring and evaluation

Finally, monitoring and evaluation of plans and programmes is undertaken to measure the performance of the multi-year programme and of yearly expenditures. They also enable comparison of performance targets and expected achievements with actual performance.

6.6 Planning for rural transport infrastructure

National and local dimensions

Rural transport infrastructure is part of a larger transportation network. There should be linkages between national and local transport planning decisions, and local decisions should not be taken independently of national considerations. Conversely, when taking national decisions, local inputs and information are necessary if the process is to be effective. All details about transport requirements, including topography, local climatic conditions, etc., are unlikely to be known fully within central government. Local residents will know of particular idiosyncrasies of the physical and social conditions of their locality, which are not known in the capital city. It is only at the local level that the nature of the demands for rural transport, as well as their costs, will be understood fully. A key aspect of planning is that there needs to be a relationship between those making public decisions and those affected by them.

Institutional arrangements

The planning process needs to recognize that the different types of rural transport infrastructure are inter-related, and therefore should be co-ordinated (Connerley and Schroeder 1996). This is complicated because current institutional arrangements generally assign

- responsibility for tertiary roads to local government, if these responsibilities have been decentralized to the local level;
- responsibility for community roads, tracks, trails and paths, is relegated, by mandate or default, to much less well-defined 'communities'.

The common practice of 'assigning' responsibility for community rural transport infrastructure to informal community groups makes the implicit assumption that these groups will assume 'ownership' of them. But true ownership, where owners take on costly responsibilities, is unlikely to be acknowledged when assigned in this unilateral manner, rather than as the result of a two-way transaction.

Objectives

Planning for rural transport infrastructure should be concerned ultimately about access, both of rural areas to the primary and secondary network, and of access within a locality (Connerley and Schroeder 1996). Planning should aim to provide infrastructure for which there is a demand, and needs to determine the types of

access that local users want, are willing to pay for, and will support through their own contribution of resources. This avoids wasting resources, particularly by designing infrastructure that exceeds requirements of local users. Determining whether demand is ‘real’ or ‘effective’ can be difficult. Although there may be an initial commitment by users to sustain infrastructure, the public nature of these produces strong temptations to avoid later responsibilities.

Past planning emphasis has been on the provision of roads. Yet, as noted in Chapter 1, benefits from roads will only be achieved if transport is forthcoming for both people and goods. Planning needs to recognize that many poor people, particularly in rural areas, do not have access to conventional forms of transport. Rural transport planning should aim to improve the provision, regulation and management of infrastructure and vehicles (motorized and non-motorized) in relation to the demands of potential end users (Howe 1999). Non-transport solutions may also be appropriate to solve some access problems, and at lower cost. For example, a programme to rehabilitate and maintain grinding mills or water sources may cost less and have larger impact than a programme of road rehabilitation (Malmberg Calvo 1998).

There is no substitute for the participation by local people in the planning process. However, participation in planning often demands substantial amounts of people’s time, even if sporadic. The poor and women, whose time is heavily subscribed, may not be able to participate even if the opportunity is offered. That risk is reinforced if participation is perceived as involving financial obligations for participants. Participatory activity should be targeted precisely and appropriately in terms of quantity, and scheduled to respect other obligations, particularly for women to plan for child care, meal preparation and their many other responsibilities (Malmberg Calvo 1998).

The currently favoured approach to participatory planning is known as ‘accessibility planning’. This is based on analytical methods developed in sciences such as agriculture and geography, including spatial mapping of variations in relevant variables, such as income, agricultural production or population density. These measures are combined with those from traffic planning, such as network analysis, origin–destination analysis and economics. The great advantage of accessibility planning is that it is based on objectively measurable data. Typical of these are

- Road condition
- Distances between relevant places relative to one another
- Objective social criteria, such as the
 - distribution of people
 - frequency of tin roofs (relative to thatched roofs – a commonly employed proxy for relative wealth in the rural third world)
 - distribution of services.

However, this multi-criteria analysis approach has the drawback that it generates a rough ranking of the potential road projects for a programme, but does not provide a measure of the benefits generated by implementing the individual projects, nor the benefits stemming from the total programme. Thus, the investor is left in the

The transport
dimension

Participation

Accessibility
planning

Box 6.1 Integrated rural accessibility planning

The integrated rural accessibility planning methodology (IRAP) identifies and ranks alternative interventions for improving rural access. At the heart of the methodology is a 10-step prioritization process designed to guide the selection of interventions to improve rural access:

- 1 Articulate the project's purpose.
- 2 Compute indicators of access to different type of services for each community (access to markets, clinics, schools, water, firewood, grinding mills).
- 3 Create a matrix resulting from these indicators and determine those communities worst off in terms of access.
- 4 Apply the 'AAAA-S' criteria: logistical Availability, financial Affordability, technical Appropriateness, socio-cultural Aceptability, and operational Sustainability.
- 5 Rank projects based on the severity of access problems and the AAAA-S criteria.
- 6 Identify optimal interventions for those locations with the worst access problems.
- 7 Assess the availability of resources, including technical expertise and monetary and other resources, such as credit facilities, materials, land and equipment.
- 8 Finalize the choice of intervention based on the resources available and the outcomes of previous steps.
- 9 Draw up final budget plans.
- 10 Devise the work programme.

Adapted from: Connerley and Schroeder 1996.

unfortunate position of knowing how best to implement the programme, but does not know if it is worthwhile pursuing. The need for investment planning to focus on transport needs rather than purely road interventions is likely to be an outcome of adopting a participatory village-based method of appraisal. The 'integrated rural accessibility planning' approach (IRAP), summarized in Box 6.1, is an example of this.

References

- Adler, H.A. (1987) 'Sector and project planning in transportation', *World Bank Staff Occasional Paper Number Four*, Washington, DC: The World Bank.
- Ben-Akiva, M. and Lerman, S.R. (1991) *Discrete Choice Analysis: Theory and Application to Travel Demand*, Cambridge, MA: The MIT Press.
- Cascetta, E. and Nguyen, S. (1988) 'A unified framework for estimating or updating origin/destination matrices from traffic counts', *Transportation Research*, **22B**(6), New York: Pergamon Press.

- Connerley, E. and Schroeder, L. (1996) 'Rural transport planning', *SSATP Working Paper No. 19*, Washington, DC: The World Bank.
- Dantzig, G.B. (1951) 'Maximization of linear function of variables subject to linear inequalities', in Koopmans, T.C. (ed.), *Activity Analysis of Production and Allocation Chpt XXI*, New York: Wiley.
- Denno, N.A. (1993) 'Network development for four step traffic models using GIS and TIGER files', *Institute of Transportation Engineers*, 63rd Annual Meeting, Arlington.
- Howe, J. (1999) 'Poverty alleviation: transport policy and socially sustainable development', *IHE Working Paper T&RE-025*, Delft: International Institute for Infrastructural, Hydraulic and Environmental Engineering.
- Malmberg Calvo, C. (1998) 'Options for managing and financing rural transport infrastructure', *World Bank Technical Paper No. 411*, Washington, DC: The World Bank.
- Overgaard, K.R. (1966) 'Traffic estimation in urban transport planning', *Report Ci 37*, Copenhagen: Acta Polytechnica Scandinavica.
- Replogle, M.A. (1991) 'Sustainable transportation strategies for third world development', *Transportation Research Record 1294*, Washington, DC: National Academy of Sciences.
- Thomson, I. (1983) 'Methods of road project evaluation in the national transportation plans of the southern cone countries of Latin America', in ICE, *Road Investment in Developing Countries*, London: Thomas Telford.
- TRL (1998) 'Guidelines for the design and operation of road management systems', *Overseas Road Note 15*, Crowthorne: Transport Research Laboratory.
- TRRL (1988) 'A guide to road project appraisal', *Overseas Road Note 5*, Crowthorne: Transport and Road Research Laboratory.

Economic appraisal

Richard Robinson

7.1 Purpose

Return on investment

Appraisal, in the widest sense, includes the analysis and assessment of social, economic, financial, institutional, technical and environmental issues related to a planned intervention. The purpose of carrying out an economic appraisal is primarily to ensure that an adequate return in terms of benefits results from making an investment, whether it be a capital project or an investment in maintenance. An additional purpose is to ensure that the investment option adopted gives the highest return in relation to the standards adopted, and to its timing. Often the appraisal includes an analysis of the net contribution that the investment will make to the society as a whole. Thus, the economic appraisal differs from that undertaken by private companies appraising commercial investments. For more detailed discussion of economic appraisal of roads, reference can be made to *Overseas Road Note 5* (TRRL Overseas Unit 1988).

Costs and benefits

Each investment decision is unique and has features that prevent analysis following an identical pattern, although the same overall approach can usually be followed. It is normal to determine the costs and benefits anticipated over a future period if no investment is made, and compare these with the costs and benefits arising from the investment. The alternative in which no investment takes place is sometimes known as the 'baseline' or 'do-nothing' case. However, it is unusual for future investment in such cases to be absolutely zero, as there is normally some kind of facility in existence, which requires some expenditure or minimum maintenance. In cases such as this, the 'do-minimum' alternative should be considered as the most realistic baseline case against which alternative investments should be compared. The choice of an appropriate 'do-minimum' case is an extremely difficult decision and has a significant influence on the size of economic return obtained. Considerable attention should therefore be given to its selection.

7.2 Preliminary consideration

Major and minor roads

Roads can be separated conveniently into two basic types: economic and social, as indicated in Box 7.1. The approach to economic appraisal will depend on which type of road is under consideration. The first part of this chapter deals with the appraisal of economic, or 'major' roads; the second part deals with the appraisal of social, or 'minor' roads. However, note that many major roads have a social as well as an economic function, and many minor roads also have an economic function.

Box 7.1 Major and minor roads

Major (economic) roads

These are those parts of the network for which the provision and management can be justified by economic cost–benefit analysis. The function of these roads is mainly traffic-related: roads are of a relatively long length, and carry relatively high volumes of traffic.

Minor (social) roads

These are roads that cannot be justified purely on economic grounds. Although they may have an economic function, their purpose has a significant social function, and social benefits are needed to justify any investment in provision or management. Most access roads in either rural or urban areas fall into this category: roads are of a relatively short length, and carry relatively low levels of motorized traffic.

Economic appraisal is carried out on a ‘project’. A project has well-defined objectives. It has both a defined start and a defined finish, and consumes resources in moving from start to finish. Thus, building a new road can be considered as a project, unlike the ‘process’ of maintenance. However, the road maintenance process can be carried out by undertaking a number of discrete projects, each designed to achieve a road maintenance objective. A road investment project has a number of stages, normally considered as the following:

Project stages

- 1 *Identification*: Identification of the investment.
- 2 *Feasibility*: Appraisal of the identified project, including preliminary considerations, alternative projects, and recommendation of preferred project; normally requires a cost–benefit analysis.
- 3 *Design and commitment*: Definition of the preferred project, including basic design, technical specifications, construction appraisal, contract strategies, and estimate of final cash cost; securing project finance.
- 4 *Implementation*: This stage includes detailed design, issue of tender invitations, assessment of tenders, placement of contracts, construction, completion and commissioning.
- 5 *Operation*: Operation and management of the new asset.
- 6 *Evaluation*: Post-evaluation of how well the project met its objectives.

These project stages are often referred to as the ‘project cycle’: this is something of a misnomer, since the steps in developing an individual project are linear rather than cyclic. Note that data collection and analysis are needed at each of these stages. This chapter is concerned with the feasibility stage.

When conducting an economic analysis of a road investment, the length of the analysis period must be determined at an early stage. Long analysis periods are normally used for roads, which can be expected to have a long life. Short analysis periods may be appropriate for small projects where the life of the investment is expected to be limited to a few years. The difficulties in forecasting traffic and other

Analysis period

eventualities for long periods into the future argue against very long analysis periods. Whatever time period is chosen, the project will usually have some residual value at the end of this period. The residual value can be approximated by the reduction in the cost of rebuilding the road, at the end of this period, to its original condition compared to the cost of building a new road. For most major road projects, an analysis period of 15 years from the date of opening is appropriate but this should be tested by the appraisal. Choosing the same analysis period as the design life of the pavement of a paved road simplifies the calculation of residual value.

Uncertainty and risk

Projects in developing and emerging countries are invariably set against a background of economic, social and political uncertainty. Thus, appraisal involves uncertainty and risk. For example, there will be a need to collect data and forecasting of trends into the future. All data collected in the field are subject to errors and some can be particularly inaccurate. By the time these data have been used to make future projections, any error can be magnified considerably. It is important to recognize that uncertainties exist and to take steps to minimize them. The impacts of uncertainty on the robustness of the appraisal need to be determined, and the appropriate methodology depends on the level of project development:

- Risk analysis is appropriate for well-defined projects – this involves formal probability analysis of the likely range of outcomes.
- Scenario analysis is appropriate where the project is exploratory – this requires the examination of a range of future possibilities that might reasonably be expected to occur, so that a range of outcomes is spanned by the analysis.
- Sensitivity analysis is appropriate for most economic appraisals – this is undertaken by varying the magnitude of the more important variables, normally one at a time, while keeping the values of the remaining variables fixed; by considering higher and lower figures than those expected, it is possible to determine how sensitive is the results of the appraisal to uncertainty.

Choice of technology Institutional issues

The choice of technology may affect the design and the cost of the project being proposed. Issues affecting this are discussed in Chapter 18.

The success of many projects will depend upon the institutional framework in which they are set. Aspects that need to be considered are the organization, staffing, training, procedures, planning, maintenance, funding and controls within the body responsible for the project. Institutional issues are discussed in the Chapters 22 and 23.

Socio-economic considerations

A number of socio-economic factors can influence the way that a project should be executed. These include:

- Social change – what will be the social consequences of the project, and what steps can be taken to deal with these?
- Employment generation – should labour or equipment-based technology be used (see Chapter 18)?
- Poverty – how will the project affect the livelihoods of poor people?
- Minorities – have any special needs of women or minority groups been taken into account in the project formulation?
- Relocation – what properties are affected by the proposed project?
- Severance – have problems of severance been taken into account?

- Road accidents – has the project been formulated considering the effects on road safety (see Chapter 4)?
- Construction consequences – has the impact of construction on the indigenous community been considered?
- Expertise and resources – does the local design organization and contracting industry have the in-house expertise and resources to mount a project of the nature and scale involved?
- Data – is the information on the local social environment, site conditions and climate likely to be reliable?

Depending on the nature of the project, there can be environmental benefits, disbenefits or, in most cases, a combination of the two. For example, construction of a by-pass may have environmental benefits to the town by-passed in terms of reduced congestion, pollution, noise and the like. This must often be balanced against environmental disbenefits associated with the by-pass cutting through virgin land on the periphery of the town. Environmental issues are discussed in Chapter 5.

Environmental issues

The key factor that determines whether or not a project is feasible, from an economic point of view, is likely to be the current and projected future level of traffic. This is irrespective of whether a major or minor road is being considered. The volume and composition of traffic needs to be known in terms of cars, light goods vehicles, trucks, buses, heavy vehicles, non-motorized vehicles, etc., to enable decisions to be taken about design and to enable economic benefits to be determined. For the structural design of paved roads (see Chapter 15), because the lighter vehicles contribute so little to pavement damage, they can be ignored, and only the number and axle loading of the heavier vehicles need to be considered. Methods of assessing traffic demand are discussed in Chapter 3, and traffic growth forecasting is discussed in Chapter 6.

Assessing traffic demand

7.3 Cost estimation

7.3.1 Estimating issues

The cost of a road project consists of the total cash expenditure incurred over the project's life. There are two main components of this:

Investment and recurrent costs

- investment cost – development (construction and upgrading), renewal, rehabilitation; and
- recurrent costs – operation and maintenance.

It should be noted that, once a capital investment has been made, this has an inevitable, on-going and never-ending consequence in terms of recurrent expenditure needs. Thus, capital expenditure decisions are optional and conditional and, therefore, political; resulting recurrent expenditures should be mandatory, unconditional and not subject to the political process. Thus, capital investments should be subject to appraisal to determine whether or not a particular investment is worthwhile. Recurrent expenditures should also be subject to appraisal to determine the optimum expenditure mix across a range of possible options. This section of the

chapter focuses on the investment costs; the maintenance component of recurrent costs is discussed later.

Accuracy requirements

The required accuracy of the cost estimate depends on the stage of the project. At the identification stage, a very crude estimate will be sufficient, but the costing needs to be carried out with increasing levels of accuracy as the project is refined through the feasibility and design and commitment stages. The cost estimate should be sufficiently accurate to

- give the decision-makers a realistic impression of the resource requirements, appropriate to the stage of project development; and
- allow for comparison of alternative project solutions.

Estimating techniques range from the broad-brush category of ‘global’ estimation to the more detailed ‘unit rate’ technique. These are summarized in *Overseas Road Note 5* (TRRL Overseas Unit 1988).

Global costs

Global cost estimation is a simple technique that relies on historical data on costs of similar projects. These are related to the overall size or capacity of the asset provided. Examples are

- cost per kilometre of road; and
- cost per square metre of bridge deck, or per cubic metre of mass concrete.

Using historical data has dangers. It is not always clear what costs are included. For example, are the costs of design and supervision included; are tax and duties included? There is also the risk of not comparing equivalent projects. For example, are the levels of quality, pavement thickness the same; are terrain and soil conditions comparable?

Unit rate

The unit rate estimation technique is based on the traditional bill of quantity approach to pricing construction work. This contains the quantities of work to be carried out, measured in accordance with an appropriate method of measurement. Historical data on unit rates or prices are selected for each item in the bill, typically taking data from recent similar contracts or published information in price books for civil engineering. Using historical data entails some risks. It is likely that the previous projects were not carried out in identical conditions, using identical construction methods and with the same duration. In addition, rates quoted by contractors in tenders are not necessarily related directly to the work items they are pricing. It is common practice for a tenderer to distribute the total funds required across the items to meet objectives such as cash flow and anticipated changes in volume of work. However, experienced estimators with good intuitive judgement and the ability to assess the realistic conditions of the work can prepare reliable cost estimates.

7.3.2 Geotechnics

Need for geotechnical information

The amount of geotechnical work that must be undertaken will depend on whether the road is new and the alignment has to be selected, or whether the project is concerned only with rehabilitation, reconstruction or upgrading. In projects for new

roads, geotechnical surveys are usually carried out to select and compare alternative routes for the road, and general ‘corridor’ studies will be required. Supporting information relating to soil conditions, earthworks, bridge sites, drainage and materials for the works are some of the important features that have to be considered. For other road improvement projects, geotechnical information is needed to determine the choice of materials that are available for use in the pavement works. Geotechnical information is particularly important in areas subject to ground instability to prevent or repair landslides (TRL 1997).

Geotechnical surveys are described in Chapter 8.

Geotechnical
surveys

7.3.3 Engineering design

The engineering design process necessary to develop costs for economic appraisal described in this book, are as follows:

- Drainage – see Chapter 10
- Geometric design – see Chapters 11 and 12
- Pavement design – see Chapter 15.

The design of structures is a specialist subject that is beyond the scope of this book. For the design of small bridges, reference can be made to *Overseas Road Note 9* (TRL 2000).

7.4 Assessment of benefits for major roads

7.4.1 Benefits considered

For major roads (see Box 7.1), the following benefits are normally considered:

- Vehicle operating costs (VOCs)
- Road maintenance benefits
- Time savings
- Reduction in accident costs
- Economic development benefits.

Different benefits will predominate for different types of road projects. VOC savings will normally be the most significant benefit of inter-urban road projects (rural trunk roads) in situations where the cost of time is relatively low. Time and accident savings will normally play a significant role in urban road projects and in situations with high traffic volumes and high costs of time. Where traffic volumes are lower, the project justification often rests on the expected development benefits. These will often be manifested as generated traffic. If development benefits and generated traffic are both being evaluated, it is important to avoid double counting. These different types of benefit are now discussed in turn.

7.4.2 Vehicle operating costs

Cost savings VOCs are normally reduced when a road is improved. Road users perceive the savings through lower expenditures in the following areas:

- Fuel consumption
- Lubricating oil consumption
- Spare parts consumption
- Vehicle maintenance labour
- Tyre consumption
- Vehicle depreciation
- Crew costs in commercial vehicles.

Road characteristics and condition VOCs are affected by the physical characteristics of the road, in terms of factors such as road width and gradient. But work now undertaken in a number of countries has also demonstrated that condition of the surface of the road also affects VOCs. The principle parameter that characterizes surface condition is 'roughness' (Sayers *et al.* 1986). Clearly, gravel and earth roads are rougher than asphalt-surfaced roads and result in higher VOCs. The relationships between VOCs and road characteristics and condition are well documented, and are described in more detail in Chapter 21.

Determination of benefits Total VOC savings will be calculated for each year of the analysis period. Traffic volumes will grow over the period, and VOC savings per vehicle will change according to the way the traffic volume and the road maintenance strategy affect road condition. Where the project results in traffic diversion, all VOCs on both the road from which the traffic has diverted and on the project road should be considered when determining benefits.

7.4.3 Road maintenance benefits

Nature of benefits An economic appraisal should include an estimation of the cost of maintaining the road, respectively, with and without (do-minimum) the proposed improvement. Depending on the nature of the investment, maintenance costs may increase (i.e. a net cost), or decrease (i.e. a net benefit). Maintenance savings can normally be expected with the following types of projects:

- paving a gravel road where traffic levels have exceeded the level of economic surface maintenance; and
- rehabilitation or renewal of a paved road that has deteriorated badly, since the improved road is less costly to maintain than the existing one.

If significant traffic diversion from other roads is expected to take place as a result of a new project, then the changing maintenance needs on the road from which the diversion took place should be considered in the estimation of benefits.

Estimating cost savings The first step in estimating maintenance cost savings is to determine the maintenance strategy that is likely to be undertaken in the future in both the project and do-minimum situations. The strategy may reflect that the nature of the road has changed, perhaps from a gravel road to a paved road. The activities expected to be

carried out in relation to time, traffic and the required service level are then costed for each analysis case. However, maintenance-costing systems implemented in organizations are notoriously unreliable. Real costs are commonly more than 100 per cent greater than those assumed by road administrations. Many costing systems only attempt to provide details of total expenditure for budgetary purposes, and it is not possible to identify in detail the activities on which expenditures have been made. Against this background, it is difficult to obtain realistic costs that can be used in the appraisal. However, in most cases, maintenance savings will be small in comparison with savings, for example, in VOCs. Nevertheless, the impact of maintenance can be large, since this may affect VOCs significantly.

7.4.4 Time-savings

Shorter road alignments and higher average speeds will lead to savings of time. The benefits of shorter journey times will accrue to passengers being carried and to the commercial vehicle fleet because higher vehicle utilization can be achieved. The time costs of commercial vehicles include standing costs, such as crew wages, vehicle depreciation and interest on capital.

Nature of savings

Travel time-savings for passengers in buses and private cars should be divided into time-savings during working hours and during non-working hours. In the absence of better data, working time can be valued at the average wage rate, plus an element to cover social overheads. The value of non-working time depends on the willingness to pay for time by those who are commuting or travelling for private purposes. Normally there are little or no data on this aspect. However, increasingly, 'stated preference' surveys are being undertaken to value non-working time. If resources permit and the population is literate, these can be a useful and cost-effective approach. It can be argued that, when unemployment is high and wages are low, the value of time is insignificant. However, the occupants of cars are normally from the highest income group of society, and are likely to value time relatively highly. Typically, in developing and emerging countries, working time is valued at around US\$1 per hour. Non-working time is then valued at a proportion of working time, typically in the range of 0–50 per cent.

Passenger time-savings

7.4.5 Reduction in accident costs

The methods of forecasting and costing road accidents are summarized in Chapter 4.

7.4.6 Economic development benefits

The economy in the vicinity of the road may benefit if a road is improved or new access is provided. It may be easier to make trips to farms or markets, or other commercial centres. There may be benefits to agricultural producers because of reduced transport costs, which enable higher prices to be obtained at the farm gate for goods that are produced. When transport costs are lowered, economic activity will be changed throughout the whole economy as the resources saved are redeployed, as producers respond to their new cost and price structure, and as consumers adjust their pattern of expenditure. As transport becomes cheaper, more people can afford

Economic development

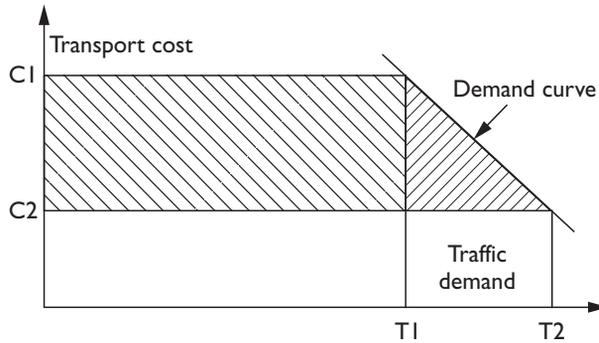


Figure 7.1 Use of a demand curve to assess generated traffic benefits.

Generated traffic

to use it. The extent to which the local economy adjacent to the road will benefit from the investment will depend on its economic potential, such as unused land and labour, and on the change in transport costs and prices.

Thus, extra benefits will accrue as a result of the generated traffic that is induced to travel on the road because of lower costs. This traffic is a measurable indicator that economic development benefits have taken place. Where the change in transport costs is relatively small, generated traffic benefits will, in most cases, represent a very small component of total benefits and can often be ignored. By contrast, when the change in transport costs is large, then generated traffic benefits are more likely to be significant. The amount of traffic that is generated will depend on the size of the unit cost reduction and on the ability of the consumer to take advantage of this cost reduction. This ability is illustrated by the demand curve shown in Figure 7.1. A cost reduction from C1 to C2 results in an increased number of trips from T1 to T2: the greater the cost reduction, the more trips that will be generated. The slope of the demand curve slope is known as the ‘elasticity of demand’, and this can normally be approximated by a straight line. Generated traffic benefits are the area under the demand curve bounded by C1, C2, T1 and T2

$$\text{Benefit} = 0.5(C1 - C2)(T2 - T1)$$

Unless better knowledge is available, it can often be assumed that this has a value of minus one (-1). This implies that a 1 per cent reduction in vehicle operating costs leads to a 1 per cent increase in traffic when considering the whole trip.

7.5 Cost-benefit analysis for major roads

7.5.1 Prices

Resource costs

For economic appraisal, the assessment is made in terms of the net contribution that the investment will make to the national economy. ‘Opportunity costs’ are used as a measure of resource rather than market prices. Costs and prices need to be adjusted to ensure that they are all measured in the same units and that they represent real resource costs to the country as a whole.

A first step in this is usually to remove the effect of inflation to enable values to be compared on the same basis over time. In most cases, it can be assumed that future inflation will affect both costs and benefits equally. Costs and prices are therefore expressed in constant monetary terms, usually for the first, or base, year of analysis.

Inflation

It is also necessary to factor costs and benefits to take account of the different economic values of investments made at different times during the analysis period. When money is invested commercially, compound interest is normally paid on the capital sum. The interest rate comprises inflation, risk and the real cost of postponing spending. Thus, money used to invest in this situation could be invested elsewhere and earn a dividend. By using capital to invest in a particular item of work or in a project, the dividend is foregone and this should be taken into account in the analysis. To do this, all future costs and benefits are discounted to convert them to present values of cost (PVC) using the formula:

Discounting

$$PVC = c_i / (1 + r/100)^i$$

where c_i is the costs or benefits incurred in year i ; r the discount rate expressed as a percentage; and i the year of analysis where, for the base year, $i = 0$.

The value of the discount rate used will have a considerable influence on the balance between the effect of investment costs, which are typically spent early in the project life, and that of benefits obtained in the future. Discounted benefits may exceed costs at one discount rate, but not at another. Since the inflation element is dealt with separately (see previous paragraph), and risk also needs separate treatment, the discount rate used may differ from market interest rates. The choice of discount rate is therefore crucial to the outcome of an appraisal in many cases.

Discount rate

Prices paid for goods and services may not reflect the real value of national resources because of distortions. These may be caused by governments fixing an unrealistic exchange rate (normally resulting in a 'black market' price for goods), because of control of imports and exports through quotas and subsidies, or because of the existence of monopolies or cartels that can 'fix' the price of certain commodities above their real resource value. The addition of taxes also distorts the real value because the taxation component represents a transfer of spending power from those purchasing the product to the government.

Price distortions

An economic cost-benefit analysis considers how best to allocate real resources to improve the rate of economic growth. Thus, the effect of all price distortions must be removed. This is known as 'shadow pricing'. For example, the official exchange rate may over-value domestic currency in relation to foreign currency, whereby imported goods appear too cheap. On the other hand, official wages are often fixed at higher levels than the opportunity cost of labour. The real resource cost of a commodity is found from:

Shadow prices

$$\text{Resource cost} = \text{financial cost} - \text{taxes} + \text{subsidies}$$

This convention adjusts financial costs to take account of the fact that taxes and subsidies are transfer payments and do not reflect the cost of resources consumed in transport.

Note that, in this respect, economic analysis using cost-benefit methods differs from the *financial cost-benefit analysis* that would be undertaken, for example, by

Market prices

the promoter of a toll road. The financial analysis would be undertaken using market prices.

7.5.2 Life cycle costing

The various costs involved in economic appraisal are inter-dependent. The adoption of higher engineering or maintenance standards normally leads to higher investment costs, but may result in lower costs to the road administration in terms of future costs of maintenance and renewal, and to lower costs to road users. Economic appraisal requires the comparison of different levels of investment at the present time compared with their respective consequential future costs. If life cycle costs are not taken into account, investment decisions become subjective and can lead to economic loss.

Investment choice

The analysis of life cycle costs for roads assumes increased complexity because costs and the relationship between costs change over time. As time passes, roads deteriorate. As levels of deterioration increase, the need for road maintenance, and hence its cost, increases. However, as roads become rougher, VOCs will increase. Travel time and accident rates may also change. Thus, road-user costs will be affected by the condition of the road surface and this, in turn, will change over time as the road is maintained. The condition of the road surface is also affected by the design and construction standards, by the traffic loading, the maintenance standards and the environment. The more vehicles that use the road, and the heavier their axle loads, the more quickly the road will wear out. The road will also wear out more quickly if it is subject to extremes of climate that may weaken its structure or cause erosion. Thus, the road standards, the environment, the vehicles and the level of maintenance all have an effect on the cost and the changes in cost experienced by the road users. This can be considered schematically as shown in Figure 7.2. These basic

Life cycle analysis

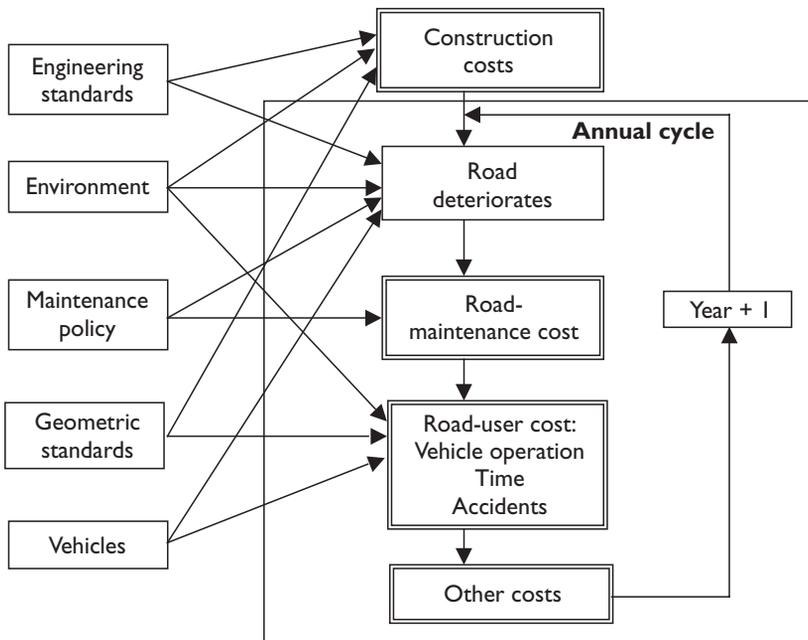


Figure 7.2 Annual cycle of cost and road deterioration.

Adapted from: Robinson (1993).

considerations under-pin the approach to modelling life cycle costs for roads, such as in HDM-4 (see Chapter 21).

7.5.3 Comparison of alternatives

7.5.3.1 Exclusivity of alternatives

As noted earlier, cost–benefit analysis is concerned with the comparison of alternatives: often ‘doing something’ compared with ‘doing minimum’. When the choice to carry out one alternative precludes the possibility of undertaking the other, this is known as ‘mutual exclusivity’, and most cost–benefit analysis is concerned with this situation. In effect, this is concerned with choosing the best way of making an investment. Sometimes, alternatives are not mutually exclusive, and the process then becomes one of ranking or prioritization.

7.5.3.2 Net present value

The ‘net present value’ (NPV) decision rule is normally used to determine whether an adequate return in terms of benefits results from making an investment. This rule may also be used to determine which investment option gives the highest return of those considered. NPV is simply the difference between the discounted benefits and costs over the analysis period.

Investment
with highest
return

$$\text{NPV} = \sum_{i=0}^{n-1} \frac{b_i - c_i}{(1 + r/100)^i}$$

where n is the analysis period in years; i the current year, with $i = 0$ in the base year; b_i the sum of all benefits in year i ; c_i is the sum of all costs in year i and r the planning discount rate expressed as a percentage.

NPV is a measure of the economic worth of an investment. The NPV can only be calculated for a predetermined discount rate, which needs to be the same for each investment being compared. The NPV should only be quoted in conjunction with the discount rate that has been used. A positive NPV indicates that the investment is justified economically at the given discount rate and, the higher the NPV, the greater will be the benefits or the lower will be the costs. Thus, the choice between investments should be based on NPV.

Interpretation
of findings

7.5.3.3 Internal rate of return

Although NPV is the correct decision rule for determining whether or not investments are worthwhile, its use is sometime impracticable. For example: an organization such as the World Bank each year needs to appraise many hundreds of projects spread throughout a number of sectors in different countries. Comparing the NPV across all potential projects to see which would offer the best investment choice is not practical. Instead, the ‘internal rate of return’ (IRR) decision rule is used. The IRR is the discount rate at which the present values of costs and benefits are equal;

Need for
alternative
method

in other words, the NPV = 0. Calculation of IRR is not as straightforward as for NPV, and is determined by solving the following equation for r :

$$\sum_{i=0}^{n-1} \frac{b_i - c_i}{(1 + r/100)^i} = 0$$

with notation as earlier. Since IRR is a relatively simple profitability ratio, it can be used in conjunction with sensitivity analysis to indicate the robustness of an investment in the face of uncertainty and risk. Note that IRR should not be used to compare mutually exclusive investments.

Interpretation of findings

The IRR gives no indication of the size of the costs or benefits of an investment, but acts as a guide to its profitability. The higher the IRR, the better the investment. If it is larger than the discount rate, then the investment is economically justified. Organizations such as the World Bank set 'cut-off' rates of return for different countries. These are based on experience of what IRR values are necessary to justify the Bank funding within a country. As such, funding an investment whose IRR is above the cut-off rate provides a convenient mechanism for selection. The cut-off IRR percentage rate is always more than the discount rate for the particular country to ensure that the NPV is positive for selected investments.

7.5.3.4 NPV/cost ratio

Efficiency of investment

One problem with the use of NPV is that, other things being equal, a large investment will normally have a larger NPV than a smaller one and, on the basis of this criterion, would always be chosen. This can cause difficulties when several potential investments are being compared under budget constraints. The problem can be overcome by considering the 'NPV/cost' ratio. This ratio represents the magnitude of the return to be expected per unit of investment and is, therefore, a measure of the efficiency of an investment, since its value is increased either by:

- increasing the size of the NPV; or
- reducing the size of the cost.

Ranking of investments

Given a constrained budget situation, the most efficient investment is that with the largest NPV/cost. This investment should be undertaken first, followed in turn by those with successively lower NPV/cost ratios, until the budget is exhausted. This approach maximizes the NPV that can be obtained from a limited budget. It enables several smaller investments, which in aggregate have a higher NPV, to be chosen over a single larger investment.

7.5.3.5 Choice of method

NPV and IRR

In most cases, the NPV and IRR will give consistent results and will identify the same investments as being worthwhile. In general, where a government is using a target, or minimum return on capital, maximizing NPV should be the criterion adopted. IRR is particularly useful when discount rates are uncertain. Normally both methods should be evaluated for an investment. Both methods can be evaluated using computer spreadsheets.

Whereas NPV and IRR are appropriate for appraising one-off investments, many decisions relating, for example, to road management require selecting those investments to make from a number that are possible, under conditions of budget constraint. In these situations, the NPV/cost decision rule is that which is most appropriate.

NPV/cost

7.6 Minor road appraisal

7.6.1 Preliminary considerations

Minor roads consist of conventional roads, built to a relatively low standard for the use of motorized vehicles, but also consist of tracks, trails and paths, used principally by non-motorized transport (NMT) – bicycles, animal transport, pedestrians, etc. Often the term ‘rural transport infrastructure’ (RTI) is used to cover this broad grouping. Whereas the aims of major roads relate broadly to providing ‘mobility’, minor roads (and RTI) tend to relate to providing ‘access’. Different types of access are described in Box 7.2. In this context, ‘basic access’ is defined as the least cost (in terms of total life cycle cost) intervention for ensuring reliable, all-season passability for the means of transport prevailing locally.

Basic access

The appraisal of minor roads differs from that for major roads for two main reasons:

Need for special considerations

- since the cost per kilometre of constructing minor roads is relatively low, the expenditure of significant amounts of time and money to determine the detailed costs and benefits of projects cannot normally be justified; and
- since the purpose of minor roads is different, the types of benefits also differ to those for major roads.

Nevertheless, economic appraisal should still be carried out for all minor road projects. In most cases, appraisal consists of a combination of screening and ranking.

Box 7.2 Access and level of service

It is useful to consider RTI, and its impact on ‘accessibility’, from the perspective of ‘level of service’. The following four levels of service, or access, can be considered:

- *No (motorized) access* – defined as no motorized access within 1–2 km of a household or village.
- *Partial access* – defined as motorized access with interruptions during substantial periods of the year (the rainy season).
- *Basic access* – defined as reliable all-season access for the prevailing means of transport, with limited periods of inaccessibility.
- *Full access* – defined as uninterrupted all-year, high quality (high speed, low roughness) access.

Source: Lebo and Schelling (2001).

Screening reduces the number of investment alternatives. This can be done, for example, through targeting disadvantaged communities based on poverty indices, or by eliminating low-priority links from the list according to agreed criteria. The remaining alternatives can be ranked according to priority.

Methodologies The following basic methodologies exist for the appraisal of minor roads and for ranking:

- Multi-criteria analysis
 - compound ranking methods
 - cost-effectiveness analysis
- Cost–benefit analysis
 - producer surplus
 - consumer surplus
 - methods using a wider range of benefits.

These are discussed in this section.

Preliminary issues to consider As for major roads, preliminary consideration will need to be given to the following:

- Choice of technology, recognizing that labour-based approaches will often be more appropriate for minor roads, as discussed in Chapter 18.
- Institutional issues for minor roads may need to recognize the particular issues relating to decentralized road administration (Robinson and Stiedl 2001).
- Socio-economic considerations, including the impact of the road investment on poverty alleviation (Gannon and Zhi Liu 1997; Ashley and Carney 1999).
- Environmental issues, as discussed in Chapter 5.
- Traffic, as discussed in Chapter 3 recognizing that traffic is particularly difficult to count and forecast on low-volume roads and that, for many minor roads, non-motorized traffic will be the dominant mode and needs to be included in the appraisal.
- Costs – ‘global cost’ estimates, as discussed earlier, should normally be used for the appraisal of minor roads.

General considerations For roads where higher than basic access standards seem justified – for example, those that have traffic levels between about 50 and 200 vehicles per day – cost–benefit analysis is recommended. In these cases, a wider range of benefits may need to be considered than for the analysis of major roads, described earlier. Multi-criteria analysis can be used where traffic is less than about 50 motorized four-wheeled (4WD) vehicles per day. Multi-criteria analysis can also be useful for screening potential investments to determine which are worth investigating in more detail.

7.6.2 Multi-criteria analysis

Basis of the methods These methods combine measurable parameters with those that are subjective. The advantage of adopting the approach is that more factors, such as road density,

accessibility and population can be included to reflect wider political and socio-economic needs. The disadvantages are:

- the addition of other factors complicates the analysis and makes it less transparent; and
- more data need to be collected for an analysis.

Compound ranking methods enable social and political factors to be considered alongside economic factors. They rank projects according to factors that are considered relevant to the investment decision. The principles of the method are:

Compound
ranking
methods

- Factors included should reflect the objectives of the road investment programme.
- Each factor is measured in its own units (e.g. number of people gaining access to services, tonnes of agricultural output, number of persons employed, etc.).
- Factors are weighted to reflect their impact on the project objectives.

These methods utilize a points-scoring method and are rather subjective. An example of an application is shown in Box 7.3. Compound ranking can lead to non-transparent results, and is recommended only if (Lebo and Schelling 2001):

- Cost criteria are included
- Criteria and relative weights are
 - few
 - relevant
 - have been determined in a participatory manner.

A sub-set of multi-criteria analysis is ‘cost-effectiveness’ analysis. This method compares the cost of interventions with their intended impacts. The World Bank allows the use of cost-effectiveness analysis in situations where benefits cannot be measured in monetary terms, or where measurement is difficult. *World Bank Technical Paper 496* (Lebo and Schelling 2001) describes how a priority index can be defined for each link based on a cost-effectiveness indicator equal to the ratio of the total life cycle cost necessary to ensure basic access, divided by the population served. Based on this, a threshold can be established below which a link should not be considered for investment. The recommended approach is to do a few sample cost–benefit analyses on a few selected links. The following cost-effectiveness formulation is suggested:

Cost-
effectiveness
analysis

$$\begin{aligned} &\text{Cost-effectiveness indicator of link}_{(i)} \\ &= (\text{Cost of upgrading link}_{(i)} \text{ to basic access standard}) \\ &\quad / (\text{Population served by link}_{(i)}) \end{aligned}$$

In a typical example (Lebo and Schelling 2001), the value of US\$50 per person served was used for the maximum amount of investment allowed per link to give a cut-off value for the investment. In practice, the threshold is not normally a point of debate since available investment budgets are normally exhausted well before what most planners agree are reasonable cost-effectiveness limits.

Example

Box 7.3 Example of compound ranking method

The indicators used for prioritization of feeder road investments are:

- *Access improvement* – the extent to which the investment improves access for different types of vehicle.
- *Length* – the length of the road.
- *People served* – the number of people dependent on the road to get economic and social services.
- *Economic/social factors* – the importance of economic and social facilities served by the road.

Access improvement is scored as follows:

Impassable to all traffic	0
Passable to 4WD vehicles when dry; impassable to 2WD vehicles	1
Passable to 4WD vehicles nearly all year; impassable to 2WD vehicles in the wet	2
Passable to 4WD vehicles all year; often closed to 2WD vehicles when wet	3
Passable to 2WD vehicles all year	4
Reliable access all year to all vehicles	5

The length in kilometres of the road is used as a factor, reflecting that a longer road benefits more people. The number of people living within the area of influence of the road is used as a factor, reflecting the population likely to benefit from the road. *Zones* are used to account for variations in agro-ecological areas, thus weighting benefits by an agricultural impact factor. Examples of zones are:

Zone:

Dry	1
Intermediate	2
Wet	4

The ranking index is given by:

$$\text{Priority index} = \text{access improvement} \times \text{length} \times \text{people served} \times \text{zonal weighting}$$

The scope of the priority index can be extended to cover other economic and social factors, but consideration of these more detailed factors may introduce complexity and spurious accuracy based on unreliable assumptions.

Adapted from: Anderson (1995).

Note
2WD – two-wheeled; 4WD – four-wheeled.

7.6.3 Cost–benefit analysis

7.6.3.1 Producer surplus

Benefits due to ‘producer surplus’ are calculated directly on the basis of the increase in farm-gate prices received by agricultural producers. The benefit is the product of the price reduction and the volume of production. Several difficulties with the application of this method have been identified (Bovill 1978). First, the net increase in agricultural production is unlikely to result simply from the road investment. Other inputs, such as irrigation or fertilizers, may well be provided, so it would be wrong to associate the total increase in production to the road investment alone. Second, there are problems of double counting. A reduction in transport costs to normal traffic is not necessarily separable from a rise in agricultural output. Third, there may be benefits to the community besides an increase in agricultural production. Increased ease of transport could affect school attendance, medical services, school and shopping trips. The data requirements for this method also pose problems both in terms of the number of parameters needed and the volume of data that may need to be collected. The World Bank has proposed simplified methods of analysis (Beenhakker and Lago 1983), but data requirements are still somewhat daunting, even for these. The approach is not generally recommended.

7.6.3.2 Consumer surplus

Consumer surplus is the direct benefit to road users from cost savings resulting from a road investment. It is the product of the number of trips and the cost saving per trip. This cost saving is known as the ‘consumer surplus’. Technically, there is only a consumer surplus if cost savings are passed on to consumers through lower fares and freight charges; otherwise they are retained by the vehicle operators as increased profit. For those projects where traffic levels are likely to be sufficient for road-user cost savings to justify funding, then the consumer surplus method should be used. The traffic level above which this method is appropriate is likely to be between 50 and 100 vehicles per day.

Road-user
benefits

VOCs tend to be correlated strongly with road roughness. If the relationship between these two parameters is known, then VOCs on the existing road or track can be estimated by measuring or estimating the roughness. For feeder road analysis, it will often be convenient to produce simple graphs of VOCs against roughness, using simple regression analysis, such as that shown in Figure 7.3. The appropriate values for roughness can be entered into the graph to give the unit vehicle operating costs on the existing and improved road.

VOC savings

7.6.4 Extending the range of benefits

Possible enhancements to cost–benefit analysis for minor roads include

Benefits

- better assessment of the costs of interrupted access or, conversely, benefits of improved traffickability;
- estimating operating cost savings of NMT;

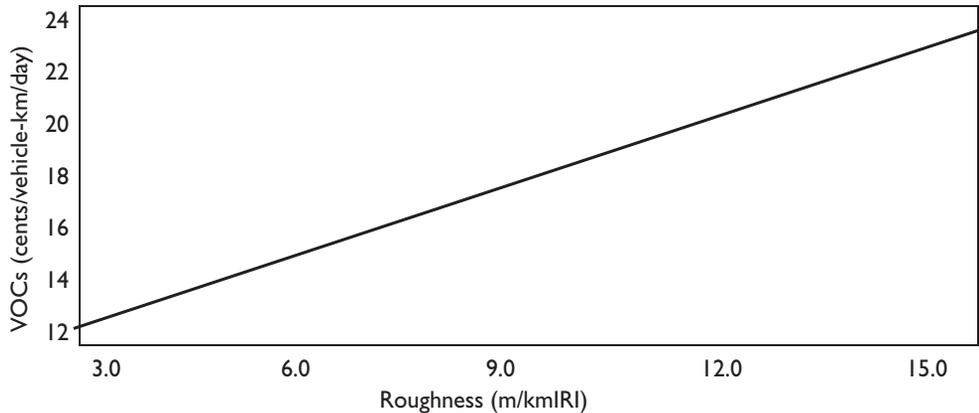


Figure 7.3 Indicative VOCs for different roughness values.

Notes

- 1 Based on the regression equation: $VOC = 11.1 + 0.83IRI$ which assumes a particular traffic mix and local circumstances and costs, with VOCs calculated using the HDM model.
- 2 With such methods, costs are specified in terms of a base date, and must be updated regularly.

- savings due to mode changes (from NMT to motorized transport);
- improved valuation of time savings, including those of pedestrians; and
- valuation of social benefits from improved access to schools and health centres.

Examples of benefit calculations are given in *Technical Paper 496* (Lebo and Schelling 2001) and by Robinson (1999).

References

- Anderson, M. (1995) 'Why are rural roads so important?', *New World Transport '95*, London: Sterling Publications, 59–67.
- Ashley, C. and Carney, D. (1999) *Sustainable Livelihoods: Lessons from early Experience*, London: Department for International Development.
- Beenhakker, H.L. and Lago, A.M. (1983) 'Economic appraisal of rural roads: simplified operational procedures', *World Bank Staff Working Paper No. 610*, Washington, DC: The World Bank.
- Bovill, D.I.N. (1978) 'Rural road appraisal methods for developing countries', *TRRL Supplementary Report 395*, Crowthorne: Transport and Road Research Laboratory.
- Gannon, C.A. and Zhi Liu (1997) 'Poverty and transport', *Transportation, Water and Urban Development Papers TWU-30*, Washington, DC: The World Bank.
- Lebo, J. and Schelling, D. (2001) 'Design and appraisal of rural transport infrastructure: ensuring basic access for rural communities', *World Bank Technical Paper No. 496*, Washington, DC: The World Bank.
- Robinson, R. (1993) 'Life cycle costing of roads', in Bull, J.W. (ed.) *Life Cycle Costing for Construction*, London: Blackie, 53–85.
- Robinson, R. (1999) 'A new approach to quantifying economic and social benefits for low-volume roads in developing countries', *Impact Assessment and Project Appraisal*, **17**(2), 147–155.

-
- Robinson, R. and Stiedl, D. (2001) 'Decentralization of road administration: case studies in Africa and Asia', *Public Administration and Development*, **21**, 53–64.
- Sayers, M.W., Gillespie, T.D. and Paterson, W.D.O. (1986). 'Guidelines for conducting and calibrating road roughness measurements', *World Bank Technical Paper No. 46*, Washington, DC: The World Bank.
- TRL (1997) 'Principles of low cost road engineering in mountainous regions, with particular reference to the Nepal Himalaya', *Overseas Road Note 16*, Crowthorne, Transport Research Laboratory on behalf of the Overseas Development Administration.
- TRL (2000) 'A design manual for small bridges', *Overseas Road Note 9*, 2nd edn, Crowthorne: Transport Research Laboratory on behalf of the Department for International Development.
- TRRL Overseas Unit (1988) 'A guide to road project appraisal', *Overseas Road Note 5*, Crowthorne: Transport and Road Research Laboratory.

Design



Motorized camel. (Photo: Polfoto)

Soil investigation

Bent Thagesen

8.1 Introduction

A soil investigation is an integral part of the location, design, and construction of roads. Soil conditions as well as topography, land use, environment problems, and political considerations, must be considered in selecting the position of the road. The soil investigation provides pertinent information about soil and rock for decisions related to the following:

Purpose

- selection of road alignment;
- the need for subgrade or embankment foundation treatment;
- investigation of slope stability in cuts and embankments;
- location and design of ditches and culverts;
- selection and design of road pavement;
- location and evaluation of suitable borrow and construction materials;
- design of foundations for bridges and other structures.

In selecting the alignment for a new road, the first step is normally to define a number of possible corridors between the end termini of the road. The next step is to select the best corridor for the proposed road and define within it one or more different alignments. These alignments are compared, and a final selection is made for design purposes. The process involves continuous searching and selecting, using increasingly more detailed knowledge at each decision-making stage.

Approach

An early phase of the soil investigation encompasses collection and examination of all existing information. This may include the identification of soil types from topographic maps, geological maps, soil maps, aerial photos, and satellite images, registration of groundwater conditions, and examination of existing road cuttings. The visual examination may be coupled with a small amount of sampling and testing. The preliminary soil investigation will help to secure a broad understanding of soil conditions and associated engineering problems that may be encountered. Furthermore, the information is of great value in planning and conducting the detailed soil investigation that is necessary at the design and construction stages.

Preliminary investigation

The detailed soil investigations may be divided into field investigations and laboratory testing. The field investigations include geophysical explorations, test pits and borings, sampling of soils and rocks, registration of soil profiles, and measurements of groundwater levels. Laboratory testing includes testing of representative samples.

Detailed investigation

8.2 Evaluation of existing information

8.2.1 Interpretation of existing maps

Topographic maps

Most countries in the world are covered by topographic maps on the scale of 1:50,000 to 1:250,000. These maps may be used as an aid to geological interpretation, to identify drainage networks, and to estimate gradients and earthworks volume. However, topographic maps may be inaccurate, and they are often out of date.

Geological maps

Most countries are covered by national geological maps of scale 1:100,000 or smaller. More detailed mapping may exist, but few developing or emerging countries have large-scale maps. Geological maps normally depict the bedrock up to the level beneath the soil. In some cases, rock types can be correlated with particular soil types but, for the road engineer, the main use of geological maps is for planning and for providing background information for the interpretation of aerial photos.

Soil maps

Soil maps are mainly produced for agricultural purposes, but only limited areas in developing and emerging countries are covered. Engineering particulars cannot be read from agricultural soil maps, but they are useful for planning purposes, because they indicate where variations in the soil types can be expected.

8.2.2 Air photo interpretation

8.2.2.1 Basis of the interpretation

Air photos

Air photos are normally stereoscopic photographs taken in preparation for land mapping. The scale is typically 1:20,000 or 1:50,000. Many developing countries have complete air photo cover, and most countries have at least 50 per cent cover. Prints of the photos are kept by the local survey department or a user department. Terrain information visible on air photos can be used for identification of most of the common bedrock types, associated residual soils, transported soils and organic soils, but detailed evaluation of ground conditions should never be made by photo interpretation alone.

Elements of investigation

Air photo interpretation for terrain evaluation is based on a systematic observation and evaluation of key elements that are studied stereoscopically (Lillesand *et al.* 1979). These are

- topography
- drainage pattern
- erosion
- photo tone
- vegetation.

Air photo interpretation is best learned through experience, and the following examples are only presented to illustrate the vast potential of the technique.

8.2.2.2 Topography

Sedimentary bedrock

Bold, massive, relatively flat-topped hills with steep hillsides indicate that the bedrock is horizontally bedded sandstone. In an arid climate, closely dissected terrain

with steep stream-side slopes is typical of horizontally bedded shale. In a humid climate, gentle sloping, softly rounded hills are the typical forms of horizontal bedded shale, while a gently rolling surface, broken by circular sinkholes, is characteristic of limestones. A series of straight or curving ridges points to tilted interbedded sedimentary rocks.

Massive, rounded, unbedded dome-like hills, often strongly jointed and with variable summit elevation, suggest that the bedrock consists of granite. A series of tongue-like flows that may overlap and interbed, often with associated cinder cones, is the typical landform of lava flows and material ejected from volcanic vents. A nearly level rock surface, often cut by major streams that form deep valleys and with columnar jointing on valley walls, indicates basalt.

Igneous
bedrock

Fan-shaped landforms occurring at the base of steep slopes, where streams discharge into an area with a more subdued relief, indicate alluvial soil deposits. Most fan materials consist of boulders, gravel, and sand with some silt that have been eroded from higher elevations and transported downslope.

Alluvial soils

An overall level relief, with a gentle downstream gradient and meandering rivers with cut-off meanders and natural levees is the typical landform of flood plains. The texture of flood plain soils varies greatly, because they have been accumulated slowly over years of shifting stream courses and over-bank flooding. Deposits from over-bank flooding are usually poorly drained silts and clays. Cut-off meanders contain standing water or poorly drained organic soils. Deposits on the inside of river bends (point bar deposits) consist mainly of sand and gravel. Natural levees contain principally sand and silt. A nearly level surface, formed where streams discharge into a lake or into the ocean, identifies a delta. Some deltas contain a great deal of sand and gravel. Other deltas have extremely variable soil conditions, as in the case of river flood plains.

Flood plains
and deltas

Mounds, hills, and ridges with a wind-swept appearance and an asymmetric cross-section are likely to be sand dunes. They are found inland from sandy beaches with onshore winds and in desert areas where material has been formed by the disintegration of sand stone.

Dunes

8.2.2.3 Drainage

Rocks and soils have characteristic drainage conditions depending on surface run-off, permeability, and internal drainage.

A dendritic drainage pattern (as in Figure 8.1), with gently curving streams, is typical of horizontally bedded sandstone. Centripetal drainage into sinkholes and very few surface streams are found in terrain with horizontally bedded shale. A 'trellis' drainage pattern (Figure 8.1), with major streams running along valley bottoms and secondary streams flowing down scarp slopes joining the major streams at right angles, indicates tilted interbedded sedimentary rocks.

Sedimentary
bedrock

A coarse-textured dendritic drainage pattern, with a tendency for streams to curve around the base of dome-like hills and secondary channels following joints, is found where the bedrock consists of granite. A rudimentary drainage pattern is associated with lava flows. Also, very few surface streams are visible where the bedrock is basalt.

Igneous
bedrock

A limited surface drainage system, but with numerous distributory (constructional) channels, is typical of alluvial fan deposits. A principal stream with few connecting

Alluvial and
other soils

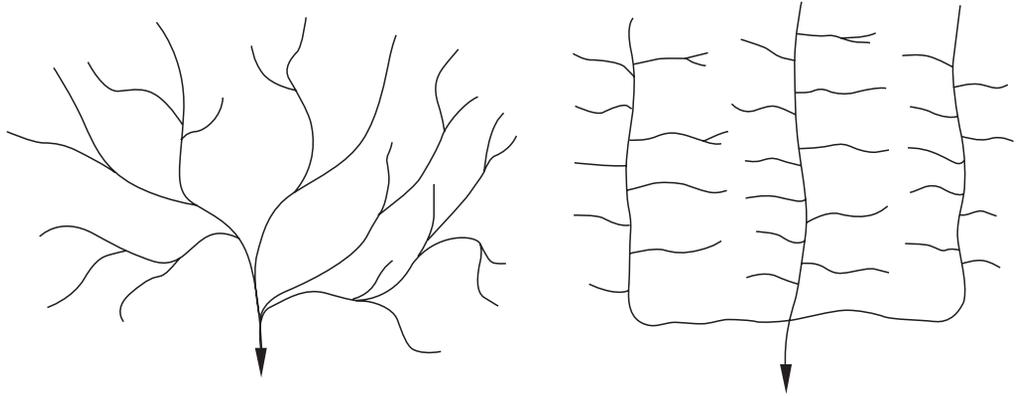


Figure 8.1 Dendritic and trellis drainage pattern.

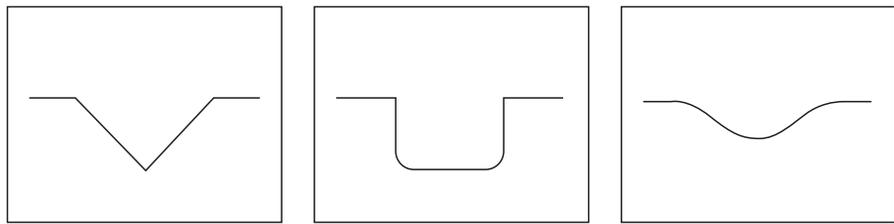


Figure 8.2 Typical gully cross-sections.

secondary streams is associated with flood plains. A major channel with branches extending in a fan-shaped pattern or in random directions is seen in deltas. Complete absence of surface drainage is characteristic of sand dunes.

8.2.2.4 Erosion

The smallest drainage features that can be seen on aerial photos are gullies. Gullies result from erosion of unconsolidated material and develop where rainfall cannot percolate adequately into the ground. Gullies collect in small rivulets and take on a shape characteristic of the material in which they are formed (see Figure 8.2). In sand and gravel, gullies tend to be short with a V-shaped cross-section. In silty soils, gullies develop with a U-shaped cross-section, and in clay soils gullies are normally long with a rounded cross-section.

8.2.2.5 Photo tone

The absolute value of the photo tone (brightness) on an air photo depends on certain terrain characteristics, but also on such factors as photographic exposure, sun angle, and cloud shadows. Because of the effect of non-terrain related factors, interpretation of photo tones must rely on relative tone value rather than absolute values. For bare soils, a light tone is typical of coarse-textured soils with low moisture content,

and low content of organic matter. A dark tone is characteristic of fine-textured soils with high moisture content and of organic soils. Coarse-textured soils will generally have a sharper borderline between light and dark tones than fine-textured soils. As the photo tone is directly dependent on the moisture content in the surface of the soil, interpretation of photo tones is difficult in arid and in very humid regions.

8.2.2.6 Vegetation

Differences in natural or cultivated vegetation often indicate differences in terrain conditions, but vegetation may also obscure differences in terrain conditions. Therefore, interpretation of vegetation should be undertaken with great caution.

8.2.3 Satellite image interpretation

During the last two decades, satellite image interpretation has come into use for geological studies and soil investigations. Satellite images are employed as a supplement to air photos, or as a substitute for air photos in cases where air photos are not available. The images of most use for soil investigations are the images produced by the Landsat, SPOT, IKONOS, and QuickBird satellites. The bulk of existing satellite images is produced by the Landsat satellites, but the number of SPOT images is growing. Satellite images are normally obtained from existing archives, as it is relatively expensive to have images specially commissioned.

Availability

The first of the Landsat series of satellites was launched by NASA in 1972. Since then, several hundred thousand images of the earth's surface have been transmitted to receiving stations operated in different countries. The sensors record information over very large areas, each scene covering 185×185 km. The first Landsat satellite had a resolution of 80 m, but later satellites have been able to produce images with a resolution of 30 m. Most Landsat images are produced by scanning in four wavelength bands. Two in the visible spectrum of green and red, and two in the reflected infrared. The satellite data are provided on films or on digital tapes. Films are normally produced at scales ranging from 1:1,000,000 to 1:250,000. In addition to black and white images of single bands, 'false' colour composites are available. The composites are generated by registering three bands onto colour film. Generally, the green band is printed in blue, the red band in green, and one infrared band in red. This combination simulates the colour rendition of infrared colour films.

Landsat

The French/Swedish SPOT satellite was launched in 1986. This satellite has the capability of collecting data similar to those from Landsat, but with somewhat better resolution. Further, the SPOT satellite is able to collect stereoscopic panchromatic images.

SPOT

Satellite imagery from the American/Japanese IKONOS satellite has become available since 2000. The resolution is 1 m for panchromatic images and 4 m for colour images. Unprocessed and ortho-rectified images can be ordered in several levels of quality.

IKONOS

The latest commercial high-resolution satellite is the QuickBird launched in 2001. The resolution is 0.6 m for panchromatic images and 2.4 m for colour images. As with the IKONOS images, a range of different products of varying accuracy is offered.

QuickBird

Use of
satellite
imagery

Because of the large scale and small resolution, satellite images are most useful for the analysis of regional features, such as terrain types, major geological boundaries and catchment areas. As with air photos, the interpretation of satellite images requires experience, and interpretation should always be checked by field investigations. Most users prefer photographic products, but the inherent digital form makes satellite data amenable to computer analysis. Quantitative techniques can be applied to interpret the digital image data automatically. Prior to displaying image data for visual analysis, enhancement techniques may be applied to accentuate the apparent contrast between features in the image.

8.2.4 Terrain classification

Background
and approach

Over the last 35 years, different terrain classification systems have been developed to provide a basis for natural resource surveys in the absence of adequate, detailed thematic maps. Terrain classification is possible because the landform is an important factor in the soil formation, and each landform in a particular area has its own geology. When classifying terrain, typical features are identified by the patterns they make in aerial photos and satellite images. These features are then used as mapping units. The main advantage of terrain classification is that the amount of soil sampling required to characterize an area is minimized. Terrain classification is particularly helpful in tropical and subtropical regions where residual soils predominate. These soils are formed over or near the rocks from where they originate.

Classification
systems

Terrain classifications already exist in many parts of the world, and these are often prepared for agricultural land-use surveys (see Figure 8.3). A system of terrain classification recommended for engineering use is described in a manual published by the Transport Research Laboratory (TRL) (Lawrance *et al.* 1993). This system is similar in concept to that used by other organizations, and land system maps made for other purposes can also be used for engineering survey. Terrain maps are commonly produced at scales of 1:250,000 to 1:1,000,000. Two levels of mapping units are normally used in the TRL terrain classification system: land systems and land facets, as in Box 8.1. Land systems and land facets are mapped on air photo mosaics of the project area. The principle in the classification method is illustrated in the block diagram in Figure 8.4.

Box 8.1 Mapping units

Land system

A land system is defined as a large area with a recurring pattern of landforms, soils, and drainage patterns. Its physical characteristics give it a distinctive, unified character, recognizable from the air.

Land facet

Land facets are the basic units that make up land systems. A land facet is a terrain unit of uniform slope, parent material, soil, and drainage pattern. The same road design and construction cost is supposed to be applicable to the whole extent of a land facet. Each land facet is subdivided into land elements often too small to be mapped.

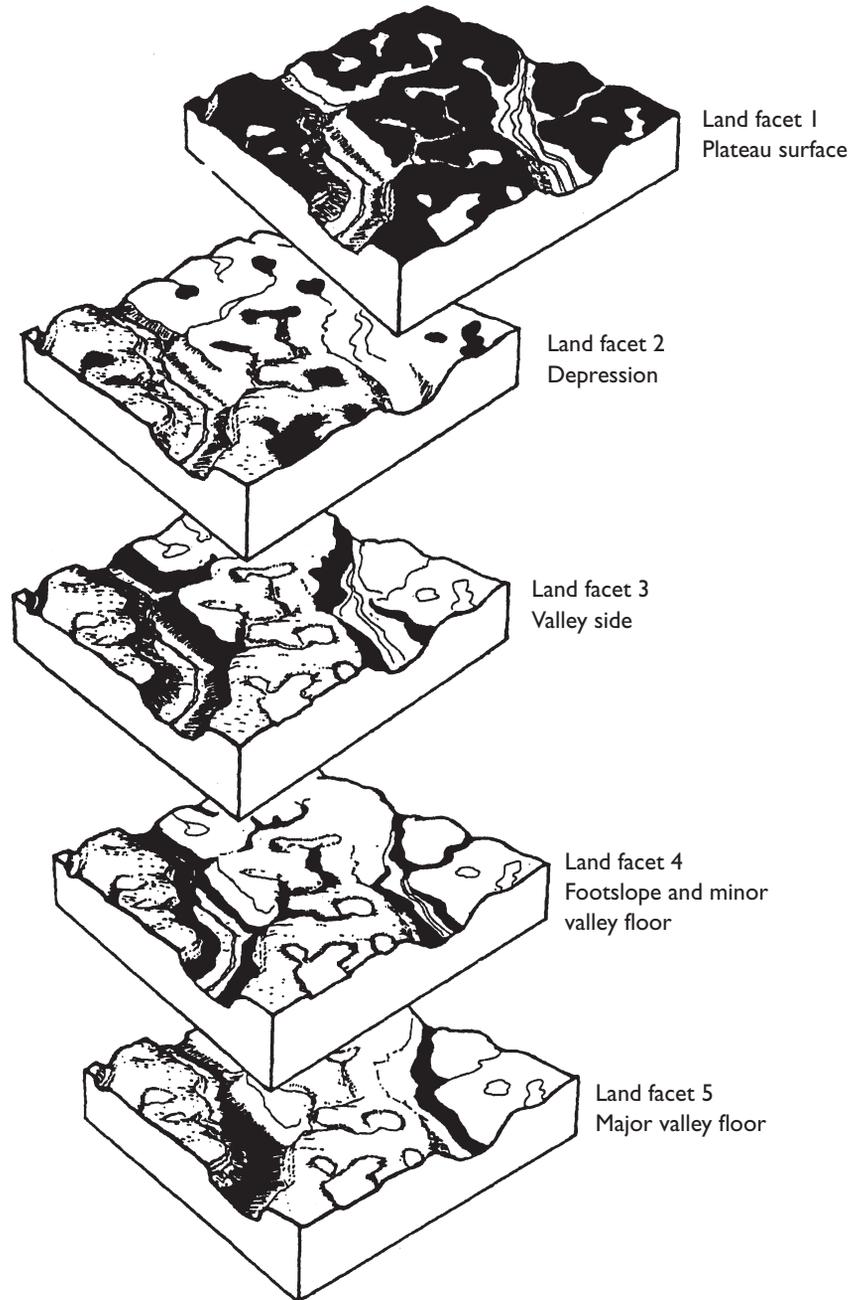


Figure 8.4 Block diagram of a land system with five land facets.
Source: Lawrance et al. (1993).

8.3 Field investigation

8.3.1 Geophysical exploration

Methods

Two geophysical methods of soil exploration are used for road purposes. These are the electrical resistivity and the seismic refraction methods.

The electrical resistivity method makes use of the varying electrical resistivity of different soils. The resistivity depends mainly on the content of clay minerals, moisture content, and the type and concentration of electrolyte in the soil-water. An increasing content of clay, water, or electrolyte causes decreasing resistivity. In performing the test, four electrodes are inserted in the surface of the soil and arranged on line symmetrical about a point (see Figure 8.5). A direct current is made to flow through the soil between the two outer electrodes, and the drop in potential is measured between the two inner electrodes. If the soil is homogeneous, the specific resistivity ρ may be calculated as

$$\rho = \frac{\pi a \left(\frac{L^2}{a^2} - 1 \right) V}{I}$$

where a is the distance between potential-electrodes; L the distance between current-electrodes; V the potential difference; and I the current intensity.

Heterogeneous soil is made up of different layers. For these, the parameter calculated is the apparent resistivity of the heterogeneous sequence down to a given depth. This depth increases with increasing distance between the current-electrodes. The electrical resistivity method may be used for soundings or for line profiling.

Soundings are carried out by taking a series of readings with increasing electrode spacing. The centre of the configuration and its orientation remain fixed. The

Electrical
resistivity

Soundings

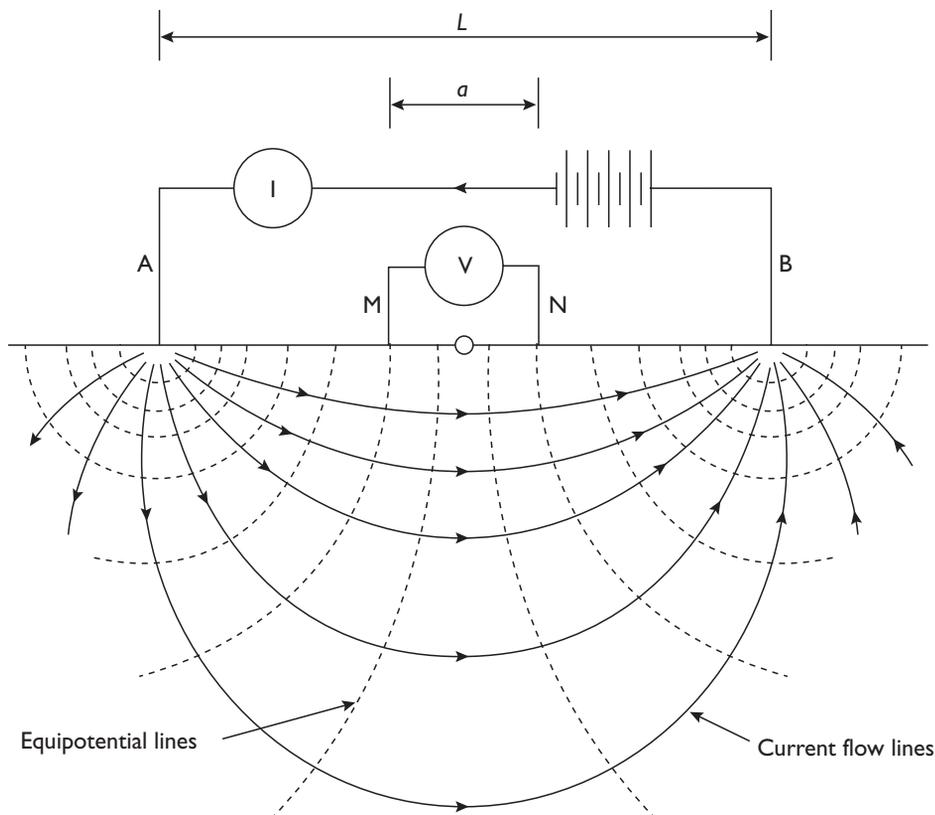


Figure 8.5 Equipotentials and current lines in a homogeneous earth.

variation of the apparent resistivity with increasing electrode spacing reflects the variation of soil conditions. The variation in resistivity may be interpreted into a soil profile showing the expected thickness and type of different soil layers in the range of measurements. Some previous knowledge of the soil types is necessary to obtain a reliable interpretation. However, an electrical sounding measures over a much bigger range than an auger boring. Since the costs of making electrical soundings are relatively small, a combination of electrical soundings and a few auger borings, for calibration, may be a very cost-effective way of soil exploration.

Line profiling

Line profiling consists of a number of measurements at points along a straight line in the terrain. The spacing of the electrodes is kept the same at all points, so all measurements correspond to the same approximate depth of penetration of the electrical field. If an area is covered with a number of line profiles, it is possible to produce a resistivity map. On this map, curves are drawn through points with the same resistivity. On a site with moist, clayey surface deposits a resistivity map may reflect the distance to stable soil strata. In sand and gravel deposits, the resistivity map will reflect the thickness of the deposits unless the measurements are disturbed by a groundwater table. In order to translate a resistivity map into layer thicknesses and examine the quality of the materials, it is necessary to make a few calibrating borings in selected locations.

Seismic refraction

The seismic refraction method relies on the principle that the velocity of sound in soils and rocks is different for different materials. A shock wave is created by detonation of a small explosive charge on the surface of the terrain. The time taken for the shock wave to reach detectors placed on a line at different distances from the source are recorded. Providing the soil is uniform to some depth, these time intervals are directly proportional to the distance from the point of detonation. If the sound velocity in a substratum is higher than in the overburden, then the time interval to more distant points is shortened because the shock wave travels through the substratum for some of the distance. By plotting travel time against distance from the point of detonation, the depth to the substratum can be calculated. The seismic refraction method is particularly useful in predicting the depth to bedrock.

8.3.2 Sampling

Methods

Soil samples near the surface are normally taken using hand tools, such as shovels and hand-operated augers. At depths greater than about 2 m, it is necessary to use power augers with appropriate samplers. Undisturbed soil samples are obtained using core-cutting devices. Rocks from prospective quarries are sampled by selection of representative boulders.

Wash drilling

Wash drilling is sometimes used for soil investigations. Wash drilling is relatively cheap and primarily used for well drilling. The equipment uses pressurized water to wash the drill rod down. Great care should be taken when using wash drilling for soil sampling. The water pressure causes extensive stirring of the soil in the vicinity of the borehole, and it may be difficult to extract undisturbed samples from the borehole even after careful cleaning of the hole.

Location

Samples of soils and rocks are extremely small compared to the quantities they represent. For example, when sampling subgrade soils, one 10-litre sample may be taken for every 100 m³ of soil. That is only 1/10,000 of the volume. Therefore, it is

of paramount importance that all samples should be as representative as possible. Sampling should be planned and targeted using topographic maps, geologic maps, air photos, etc. Geophysical investigations are also a useful guide to selecting suitable locations of test pits and borings.

The depth of test pits and borings for road beds should be at least one metre below the proposed subgrade elevation. Where soft soils are encountered, it may be necessary to increase the depth down to a denser stratum. Borings for structures or embankments should extend below the level of significant influence of the proposed load. Where drainage may be influenced by either pervious water-bearing materials, or impervious materials that can block internal drainage, borings should be sufficient in number and depth to outline the required quantities of material.

In order to secure representative samples, it is advisable to take a greater number of samples in the field than can possibly be tested in the laboratory. The field samples are then classified visually, and typical samples are selected for testing.

Depth

Number of samples

8.4 Laboratory testing

8.4.1 Approach

To a large extent, the same type of test methods are used all over the world. In specifications for testing of soils and aggregates reference is frequently made to British Standards (BSI 1984–95, 1990), standards issued by the American Society for Testing Materials (ASTM 2002), and standards specific to road engineering issued by the American Association of State Road and Transportation Officials (AASHTO 2002). However, when actually performing tests, it is of utmost importance that the specified standards be followed precisely, as small differences in the testing procedure may have a major impact on the test result.

The size and type of samples required are dependent upon the test to be performed. Normally, it is necessary to reduce the field samples into one or more smaller laboratory samples. Reduction of disturbed samples should always be carried out by quartering (see Figure 8.6), or by use of a sample splitter.

Test methods

Sample reduction

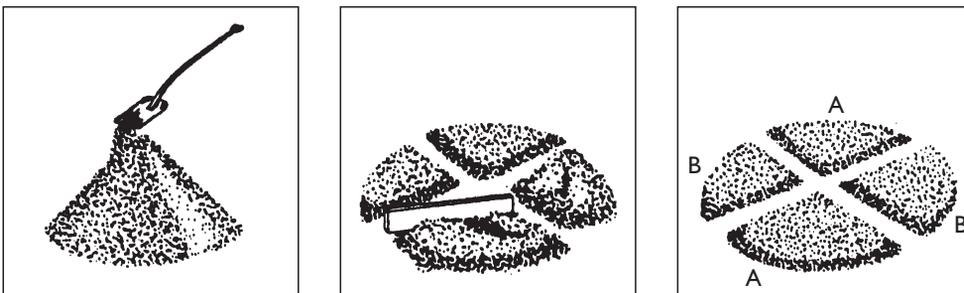


Figure 8.6 Quartering.

8.4.2 Testing of disturbed soil

8.4.2.1 Particle-size distribution

Importance	The distribution of particle sizes in soils is important in road engineering since the value of many properties, such as internal friction, voids content, wear resistance, and permeability, depend on the gradation. The distribution of particle sizes larger than 75 μm is determined by sieving a sample through a number of standard sieves.
Test method	According to most test standards, the soil sample should be placed on a 75 μm sieve and washed free of all fine materials before the sample is dried and sieved. Alternatively, the soil should be washed through successive sieves with water. However, many laboratories practice dry sieving of unwashed samples. In some tropical soils, the fines tend to stick to the coarse particles. Therefore, dry sieving should only be allowed if it has been shown that the same results are obtained as with wet sieving or sieving of washed samples. The sieves may be shaken by hand, but a mechanical shaker ensures more reliable results. All sieves wear out with use. The small sieve sizes are particularly vulnerable, and they should be inspected regularly.
Analysis	The results of a sieve analysis are normally displayed in a graph. The sieve sizes are plotted on a logarithmic scale as the abscissa. The proportions, by mass, of the soil sample passing the corresponding sieves are plotted on an arithmetic scale as the ordinate. Some typical particle-size distribution curves or sieve curves are shown in Figure 8.7. A well-graded soil is one with a gently sloping sieve curve, indicating that the soil contains a wide range of particle sizes. A uniformly graded soil is one with predominance of single sized particles. A gap-graded soil has one size range of particles missing.
Coefficient of uniformity	The coefficient of uniformity is sometimes used as a single numerical expression of particle-size distribution for purposes of succinct communication. The coefficient is defined as the ratio of the sieve size through which 60 per cent of the material passes to that of the sieve size through which 10 per cent passes.

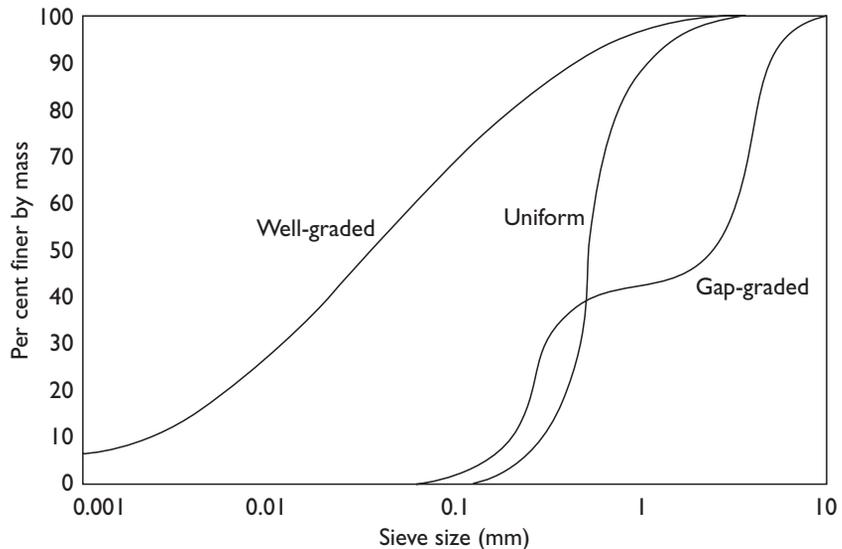


Figure 8.7 Particle-size distribution curves.

8.4.2.2 *Moisture content*

The engineering properties of a soil, such as the strength and deformation characteristics, depend to a very large degree on the amount of voids and water in the soil. The moisture content is defined as the mass of water contained in a soil sample compared with the oven-dry mass of the sample. It is customarily expressed as a percentage, although the decimal fraction is used in most computations.

8.4.2.3 *Specific gravity*

The specific gravity of a soil is used in the equations expressing the phase relation of air, water, and solids in a given volume of material. The specific gravity of a soil is calculated as the ratio between mass and volume of the solid particles of a sample. The volume of the particles is determined by placing the sample in a volumetric bottle (pycnometer) filled with water and measuring the volume of displaced water.

8.4.2.4 *Plasticity*

The plasticity limits are used to estimate the engineering behaviour of clayey soils and form an integral part of several engineering classification systems. The plasticity limits include the liquid limit (LL) and the plastic limit (PL), and they are determined by arbitrary tests on the fine soil fraction passing the 42 μm sieve.

The LL is determined by performing trials in which a sample is spread in a metal cup and divided in two by a grooving tool. The LL is defined as the moisture content of the soil that allows the divided sample to flow together, when the cup is dropped 25 times on to a hard rubber base.

The PL is determined by alternately pressing together and rolling a small portion of soil into a thin thread causing reduction of the moisture content. The PL is defined as the moisture content of the soil when the thread crumbles.

The difference between the LL and the PL is called the plasticity index (PI).

In some countries, the PI and the percentage of fines are combined into a plasticity product. This approach has considerable merit, as the effective contribution of the plasticity of the fines to the performance of the whole material depends on the proportion of fines.

Test specimens for plasticity testing are often prepared by mixing oven-dried material with water. However, the plasticity of certain tropical soils decreases when the soil is dried before testing. Therefore, it may be necessary to prevent drying of the material and use 'wet preparation' of the test specimens. Other soils may contain weak aggregations that break down under intense mixing with an attendant increase in plasticity. This also calls for wet preparation, as this procedure involves less mixing than dry preparation.

8.4.2.5 *Free swell*

Expansive clays are a problem in many regions in the tropics. A simple test can be used to verify swelling tendencies. A measured volume of dry, pulverized soil is poured into a graduated glass containing water. After the soil comes to rest at the

Plasticity limits

Liquid limit

Plastic limit

Plasticity index
Plasticity
productPreparation of
samples

bottom of the cylinder, the expanded volume is measured. The free swell is calculated as the increase of volume as a percentage of the initial volume.

8.4.2.6 Density

- Field density** The field density (in-place density) has a great influence on the bearing capacity and the potential for settlement. Soil compaction is therefore an important component of road construction because it increases the density. Measurements of field density are made during soil investigation, but most measurements are taken to assist with compaction control during construction. Several methods are used for determining field density.
- Drive cylinder** The simplest method is the drive cylinder or core cutter method where a fixed volume of soil is removed by driving a thin-walled cylinder down into the soil (see Figure 8.8). The sample is brought to the laboratory, and the dry mass determined. The dry density is calculated by dividing the oven-dried mass of the soil specimen by its volume. The method is not applicable to friable soils or soils containing coarse material.
- Sand cone** The sand cone or sand replacement method is widely used to determine the density of compacted soils. A sample is removed by hand excavation of a hole in the soil. The in situ volume of the sample is then determined by measuring the volume of dry, free-flowing sand necessary to fill the hole. A special cone is used to pour the sand into the hole (see Figure 8.9). The dry mass of the sample is determined in the laboratory. The method is not recommended for soils that are soft, friable or in a saturated condition. The method is rather time-consuming.
- Nuclear density gauge** The nuclear density gauge utilizes gamma rays to measure the wet density (total density) and neutron rays to measure moisture content. The results are available on the spot in a matter of minutes after the test has been completed. Strictly speaking, the method is not a laboratory test. It is necessary to calibrate the results with direct measurements of density and moisture content of each soil type encountered on site. The nuclear density gauge is potentially hazardous, and care is required in the

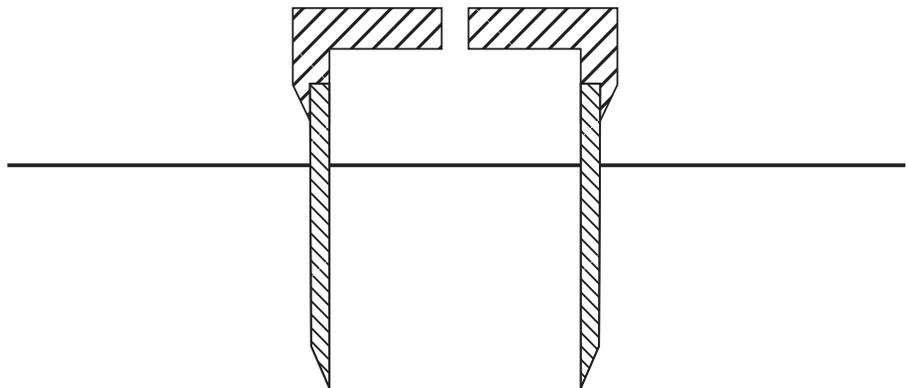


Figure 8.8 Drive cylinder.

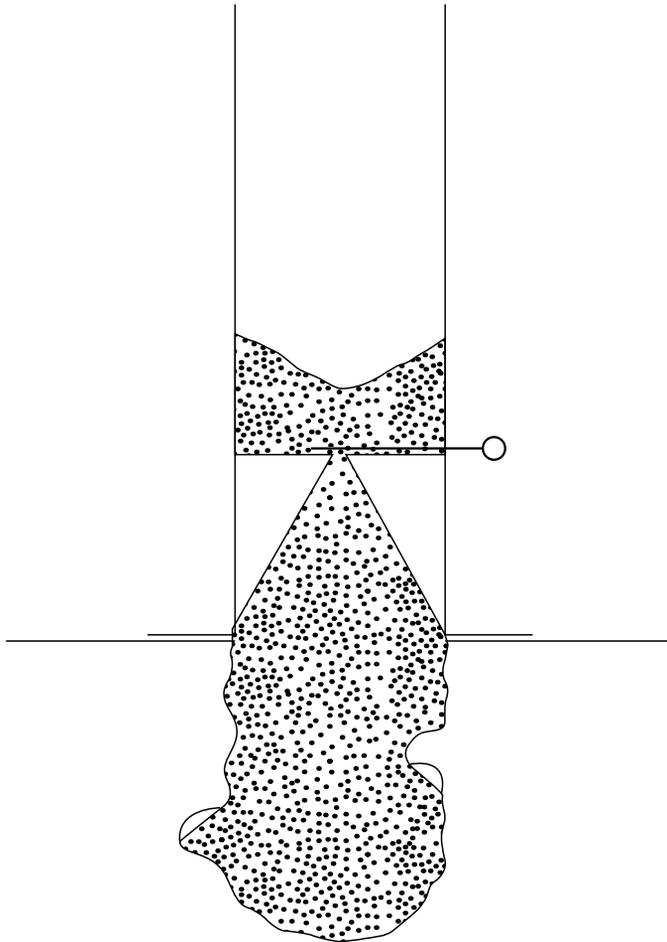


Figure 8.9 Sand cone.

handling of the equipment. It is costly and mainly used for compaction control on large equipment-based road works.

8.4.2.7 Compaction

The level of compaction to be achieved in the field during construction is normally specified as a percentage of the maximum dry density obtained in a compaction test in the laboratory. The traditional laboratory tests are the 'standard' and the 'modified' AASHTO compaction or the 'light' and 'heavy' British Standard (BS) compaction. They are also known as standard and modified 'proctor tests' after the person who invented the laboratory compaction tests.

Type of tests

A sample of the soil is compacted in a cylindrical metal mould having diameter 100 or 150 mm, with an approximate volume of 1 or 2 litres, respectively. Compaction is achieved by use of a falling hammer. For modified compaction, a heavier hammer plus more and longer hammer drops are used than for standard compaction. The

Method

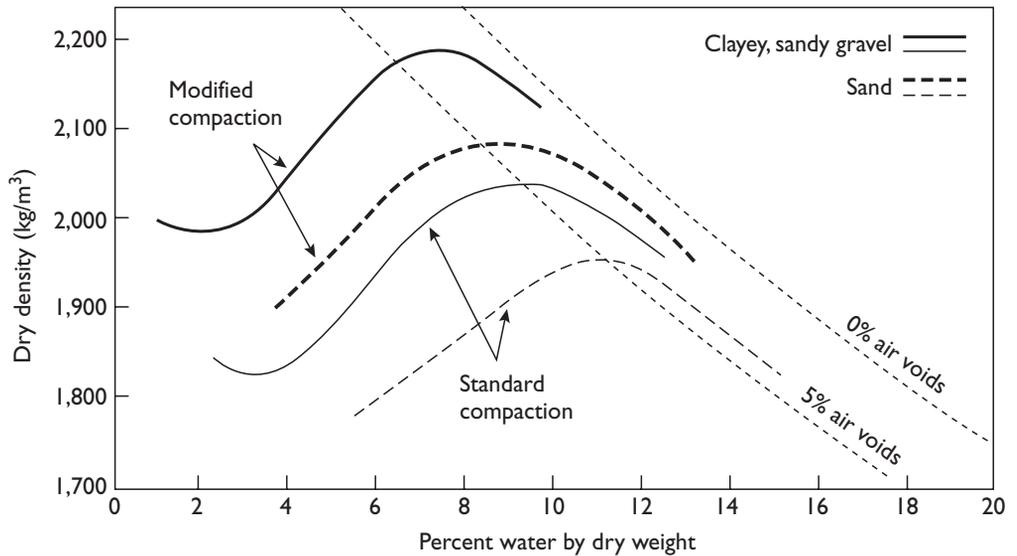


Figure 8.10 Example of compaction curves.

Source: Millard (1993).

compactive effort is designed to be equivalent to a medium-sized roller in the field. Modified compaction represents more accurately what can be successfully obtained in the field, but standard values are widely used, particularly for subgrades. The compaction test is repeated for a range of different moisture contents in the soil. For each test, the achieved dry density is recorded. Corresponding values of moisture content and dry density are then plotted in a diagram, and the points connected with a smooth curve (see Figure 8.10).

Normally, the compaction curve has an obvious peak, indicating that there is an optimum moisture content at which the maximum dry density will be achieved for a particular compaction effort. However, the optimum moisture content is not a constant for a particular material. The optimum moisture content is dependent on compaction effort and compaction method as well as on soil type. For the same soil, a greater compactive effort results in a higher maximum dry density at a lower optimum moisture content. For the same compactive effort, gravelly soils have a higher maximum dry density and a lower optimum moisture content than clayey soils.

During construction of major roads, compaction in the field is achieved by rolling. In order to obtain a high dry density, it is common practice to try to bring the moisture in the soil close to the optimum corresponding to the applied compaction effort and method. If a soil layer is too wet, it is left exposed to wind and sun for some time and allowed to dry. If the soil layer is too dry, the moisture content is increased by spraying of water. In dry tropical regions it may be difficult to obtain the quantity of water necessary to increase the moisture content in a soil layer to the optimum. Under those circumstances it may be favourable to compact the soils in the dry state. For fine-grained soils a small moisture content creates cohesion between the

Variation in optimum

Field approach

particles that impedes the compaction. This is illustrated as a minimum in the beginning of the compaction curve, as in Figure 8.10.

Certain tropical soils contain weak aggregations that break down under intense remoulding. For this type of soil, fresh material should be used for each point on the compaction curve. As an alternative, the laboratory testing may be replaced by full-scale testing in the field. The tendency to degradation may be estimated by comparing the results of a sieve analysis before and after a compaction test.

Because of the limited dimensions of the test equipment, the proctor tests are not suitable for soils containing substantial quantities of coarse particles. When using the 1-litre mould, material retained on the 4 mm sieve is normally discarded. When the 2-litre mould is used, the maximum particle size is increased to 16 or 20 mm. Compaction tests on soils containing many large particles are discussed later in the section dealing with testing of aggregates.

Material
degradation

Suitability

8.4.2.8 California bearing ratio

The California bearing ratio (CBR) test is the most common test for evaluating the bearing capacity of subgrade soils. It measures the force needed to cause a plunger to penetrate 2.5 or 5 mm into a soil sample compacted into a 2-litre cylindrical mould with a diameter of 150 mm, as shown in Figure 8.11. The measured force is taken as a percentage of a standard force.

Method

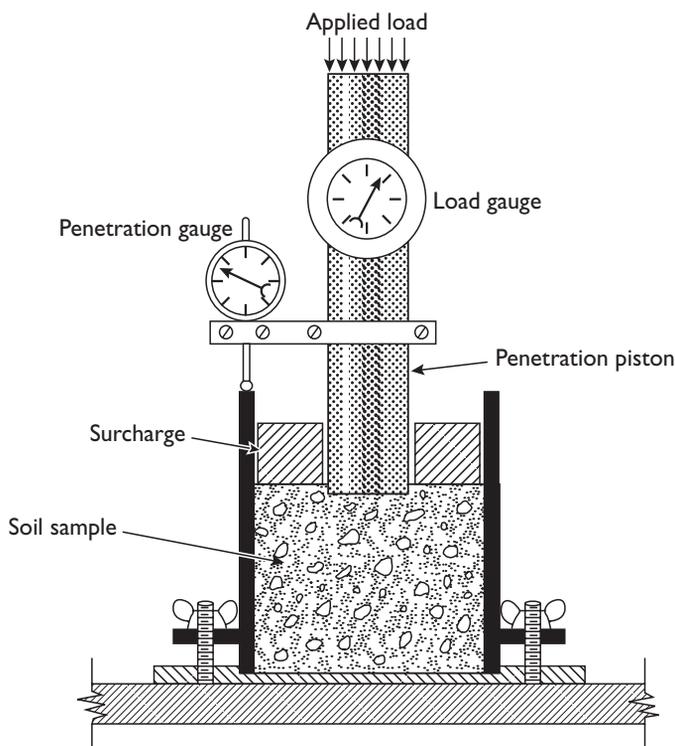


Figure 8.11 Equipment for CBR testing.

Moisture content	One of the main problems with CBR testing is deciding what moisture content to use in the soil sample. In many countries, the standard procedure is to compact the soil at optimum moisture content and then soak the sample in water for four days. However, four days of soaking would be very conservative in arid and semi-arid areas. On the other hand, in wet areas, four days of soaking may not be enough with clayey soils of low permeability. Normally, it is more useful to derive the CBR from tests on samples where the moisture content has been adjusted to the critical level likely to occur in the field. Selection of proper moisture content is discussed in the Chapter 15.
Field testing	CBR tests can be carried out in the field. Test equipment is attached to a truck, and the truck used as a dolly. For practical reasons, in situ equipment is normally only used on existing roads. The base, sub-base, and subgrade are exposed by digging holes through the pavement.
Reproducibility	The CBR test is of poor reproducibility, particularly with granular soils.

8.4.2.9 Dynamic cone penetrometer

The dynamic cone penetrometer (DCP) is a quick and cheap alternative to in situ CBR tests. A 30° steel cone is forced into the soil by use of a drop hammer, and the penetration is measured in millimetres per blow. Empirical relations between penetration and CBR may be derived.

8.4.3 Testing of undisturbed soil

Undisturbed samples are normally used for consolidation testing and triaxial compression testing. The consolidation test is employed to estimate the settlement of soils under the load of an embankment or other structure. The triaxial compression test is used to examine the structural strength of soils as foundations for structures or in detailed studies of slope stability problems. Consolidation testing and triaxial testing in particular require complicated laboratory equipment and expertise.

8.4.4 Testing of aggregates

Particle size analysis	Particle size analysis on aggregates is carried out using the same procedure as described for soils. Circular sieves with a frame diameter of 200 mm are normally used for analysis of soils and fine aggregate. However, for analysis of coarse aggregate it is useful to employ sieves with a frame diameter of 300 mm or more, because bigger samples are needed to obtain representative results. An important use of the sieve curve is for estimating the volume occupied by different fractions of the soil. In some types of natural gravel, particularly laterite, there may be a significant difference between the specific gravity of the coarse and the fine particles. For these types of soils, it may be useful to convert mass proportions to volume proportions when plotting the sieve curve.
------------------------	--

The specific gravity of aggregates is used for converting mass to volume. Volume calculations of aggregates are primarily used in connection with mix design for cement and asphaltic concrete. The test procedure is similar to that described for soils, except that bigger samples and bigger pycnometers are needed for coarse aggregate. Instead of using a volumetric bottle, the volume of the sample may be determined by placing the sample in a wire basket and weighing it before and after immersing in water.

Specific gravity

High porosity of aggregates may be a sign of low mechanical strength. Furthermore, aggregates with high porosity may be difficult and costly to dry during processing of asphalt hot mix. The porosity is estimated by measuring the water absorption. This is determined by immersing a dry sample in water for 24 h. The surfaces of the particles are then dried by rolling the sample gently in a dry cloth. The water absorption is calculated as the difference in mass between the saturated, surface-dry sample and the dry sample as a percentage of the mass of the dry sample.

Water absorption

The sand equivalent test is useful for evaluating the plastic properties of the sand fraction of aggregates. A volume of damp aggregate passing the 4.75 mm sieve is measured. The material should not be dried before testing as this may change its properties. The sample and a quantity of flocculating (calcium chloride) solution are poured into a graduated glass and agitated. After a prescribed sedimentation period, the height of sand and the height of flocculated clay are determined. The sand equivalent (SE) is the height of sand as a percentage of the total height of sand and flocculated clay in the glass.

Sand equivalent

Field density tests are used in construction to evaluate the compaction achieved in aggregate base and sub-base. Common methods are sand cone and nuclear density gauge, as described earlier. However, measurements of field density are not very precise when dealing with coarse material. In some countries, this has led to the use of method-specifications for compaction instead of end-product specifications (see Chapter 16).

Field density

AASHTO compaction tests are not suitable for testing materials with a high content of coarse particles. Vibro-compaction is a more useful method for coarse materials. A sample of the aggregate is placed in a large cylindrical metal mould having a diameter of 280 mm, and an approximate volume of 14 litres. Compaction is achieved by applying a surcharge to the surface of the soil, and then vertically vibrating the mould, soil, and surcharge. The equipment is heavy and requires a laboratory provided with lifting devices. Only a few soils laboratories are able to perform vibro-compaction. The best alternative is to make full-scale compaction tests in the field using appropriate rollers.

Vibro-compaction

The CBR test is unsuitable for testing of crushed stone and coarse gravel, because of the need for removing particles bigger than 20 mm. For design purposes, the CBR is sometimes estimated based, not on testing, but on previous experience combined with evaluation of the shape of the particle-size distribution curve.

California bearing ratio

The particle shape influences the compaction and strength characteristics of aggregate mixtures. Cubic particles are less workable but more stable than flaky and elongated particles. The particle shape test is performed on coarse particles, for example, particles retained on the 6 mm sieve. Each particle is measured using a length gauge. Particles with a smallest dimension less than 0.6 of their mean size

Particle shape

are classified as 'flaky'. Particles with a largest dimension more than 1.8 times their mean size are classified as 'elongated'. The mean size is defined as the mean of the two sieve sizes between which the particle is retained in a sieve analysis. The percentage by mass of flaky particles in a sample is called the 'flakiness index'.

Soundness

The soundness test is used as part of the materials survey and design process to estimate the soundness of aggregate when subjected to weathering. The test subjects samples to repeated immersion in saturated solutions of sodium or magnesium sulphate, followed by drying. The internal expansive force, derived from the rehydration of the salt upon re-immersion, simulates the weathering action. The sample is sieved before and after the test, and the percentage of loss for each fraction is calculated. The precision of the test is poor, and it is not usually considered for outright rejection of aggregates without confirmation by other tests.

Sulphate content

The existence of sulphates in concrete aggregates is extremely destructive. Materials containing sulphate may occur in deserts and semi-deserts. Aggregates are tested for sulphate by washing a small sample with hydrochloric acid. The hydrochloric acid converts the sulphate to sulphuric acid. Then the sample is filtered, and a barium chloride solution is added to the filtrate. The barium chloride reacts with the sulphuric acid forming insoluble barium sulphate. The sulphate content is calculated from the mass of dry precipitate.

Los angeles abrasion

The los angeles abrasion test gives an indication of the resistance to abrasion in combination with the impact strength of coarse aggregates. The test is used for selecting the most suitable aggregate sources for quarrying. A sample is loaded together with a number of steel balls into a steel drum, which revolves on a horizontal axis, as shown in Figure 8.12. The los angeles abrasion value is the percentages of fines passing the 1.7 mm sieve after a specified number of revolutions of the drum.

Aggregate impact

The aggregate impact test is a cheap alternative to the los angeles abrasion test. The sample is placed in a cylindrical steel mould and exposed to 15 blows of a falling hammer. The percentages by mass of the resulting fines passing the 2.36 mm sieve is called the aggregate impact value (AIV). The AIV is normally about 105 per cent of the so-called aggregate crushing value (ACV), which is obtained by a similar test using a compressing machine instead of a falling hammer.

Polished stone

The resistance to polishing is an important factor when selecting suitable sources for aggregates to be used in bituminous surfacing. However, polishing tests are relatively complicated and only undertaken in special laboratories.

Organic impurities

The test for organic impurities is intended to identify the presence of deleterious organic compounds in natural sands to be used in cement concrete. A quantity of sand and a volume of NaOH solution are added to a glass bottle and allowed to stand for a prescribed period. If the liquid turns darker than a standard colour, it indicates that the sand may contain deleterious organic matters.

Affinity for bitumen

Different test methods exist for determining the resistance to stripping, that is, the separation of a bitumen film from the aggregate resulting from the action of water. However, none of the methods are very reliable. In the simplest method, the resistance to stripping is determined by immersing an uncompacted sample of bitumen-coated aggregate in water. At the end of a soaking period, the percentage of surface

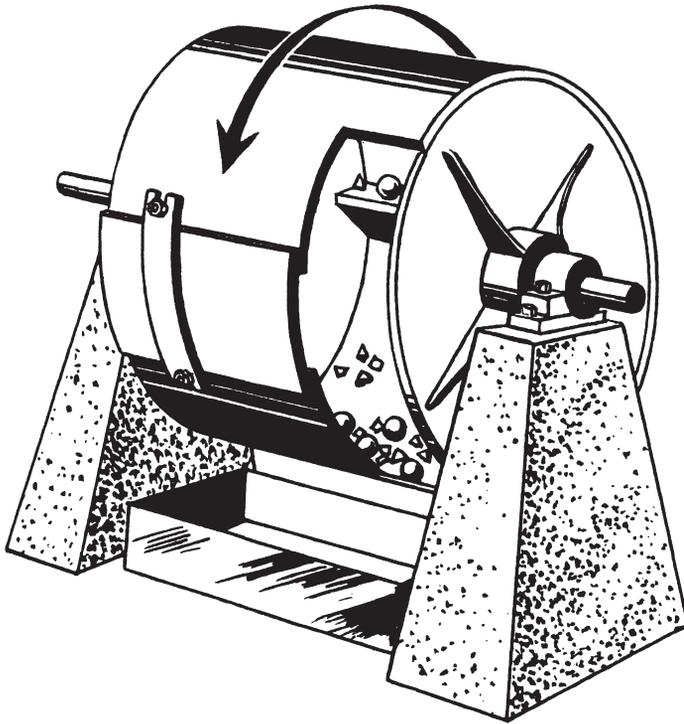


Figure 8.12 Los angeles abrasion machine.

Source: Asphalt Institute (1989).

area of the aggregate on which the bitumen film is retained is estimated visually. In the immersion-compression test, the resistance to stripping is measured indirectly. Bitumen-coated aggregate is compacted into duplicate sets of tests cylinders. The compression strength of the cylinders is measured, one set dry and the other after immersion in water. The difference in strength serves as a measure of the effect of moisture.

8.5 Soil classification

Soil classification is a way of systematically categorizing soils according to their probable engineering characteristics. The classification of a soil is based on its particle-size distribution and, if the soil is fine-grained, on its plasticity (LL and PI). The most widely used classification systems used in road engineering are the unified soil classification system, shown in Figure 8.13, and the AASHTO classification system, shown in Table 8.1. Soil classification should only be regarded as a means of obtaining a general idea of soil behaviour and it should never be used as a substitute for detailed investigations of soil properties.

Soil
classification

Criteria for assigning group symbols and group names using laboratory tests^a

Soil classification				
Group symbol	Group name ^b			
Coarse-grained soils More than 50% retained on No. 200 sieve	Clean gravels Less than 5% fines ^c	Cu ≥ 4 and I ≤ Cc ≤ 3 ^e Cu < 4 and/or I > Cc > 3 ^e	GW GP	Well-graded gravel ^f Poorly graded gravel ^f
	Gravels More than 50% of coarse fraction retained on No. 4 sieve	Gravels with fines More than 12% fines ^c	GM GC	Silty gravel ^{f,g,h} Clayey gravel ^{f,g,h}
Fine-grained soils 50% or more of coarse fraction passes No. 4 sieve	Clean sands Less than 5% fines ^d	Cu ≥ 6 and I ≤ Cc ≤ 3 ^e Cu < 6 and/or I > Cc > 3 ^e	SW SP	Well-graded sand ⁱ Poorly graded sand ⁱ
	Sands 50% or more of coarse fraction passes No. 4 sieve	Sands with fines More than 12% fines ^d	SM SC	Silty sand ^{f,g,h,i} Clayey sand ^{f,g,h,i}
Silt and clays Liquid limit less than 50	Inorganic	PI > 7 and plots on or above 'A' line ^j PI < 4 or plots below 'A' line ^j	CL ML	Lean clay ^{k,l,m} Silt ^{k,l,m}
	Organic	Liquid limit – oven dried < 0.75 Liquid limit – not dried	OL	Organic clay ^{k,l,m,n} Organic silt ^{k,l,m,o}
Silt and clays Liquid limit 50 or more	Inorganic	PI plots on or above 'A' line PI plots below 'A' line	CH MH	Fat clay ^{k,l,m} Elastic silt ^{k,l,m}
	Organic	Liquid limit – oven dried < 0.75 Liquid limit – not dried	OH	Organic clay ^{k,l,m,p} Organic silt ^{k,l,m,q}
Highly organic soils	Primarily organic matter, dark in colour and organic odour		PT	Peat

a Based on the material passing the 3-in. (75-mm) sieve.

b If field sample contained cobbles or boulders or both add 'with cobbles or boulders, or both' to group name.

c Gravels with 5–12% fines require dual symbols:

GW–GM well-graded gravel with silt

GW–GC well-graded gravel with clay

GP–GM poorly graded gravel with silt

GP–GC poorly graded gravel with clay

d Sands with 5–12% fines require dual symbols:

SW–SM well-graded sand with silt

SW–SC well-graded sand with clay

SP–SM poorly graded sand with silt

SP–SC poorly graded sand with clay.

$$e \quad Cu = D_{60}/D_{10} \quad Cc = \frac{(D_{30})^2}{D_{10} \times D_{60}}$$

f If soil contains ≥ 15% sand, add 'with sand' to group name.

g If fines classify as CL–ML, use dual symbol GC–GM or SC–SM.

h If fines are organic, add 'with organic fines' to group name.

i If soil contains ≥ 15% gravel, add 'with gravel' to group name.

j If Atterberg limits plot in hatched area, soil is a CL–ML, silty clay.

k If soil contains 15–29% plus No. 200, add 'with sand' or 'with gravel', whichever is predominant.

l If soil contains ≥ 30% plus No. 200, predominantly sand, add 'sandy' to group name.

m If soil contains ≥ 30% plus No. 200, predominantly gravel, add 'gravelly' to group name.

n PI ≥ 4 and plots on or above 'A' line.

o PI < 4 or plots below 'A' line.

p PI plots on or above 'A' line.

q PI plots below 'A' line.

Figure 8.13 Unified classification system.

Source: ASTM 2002.

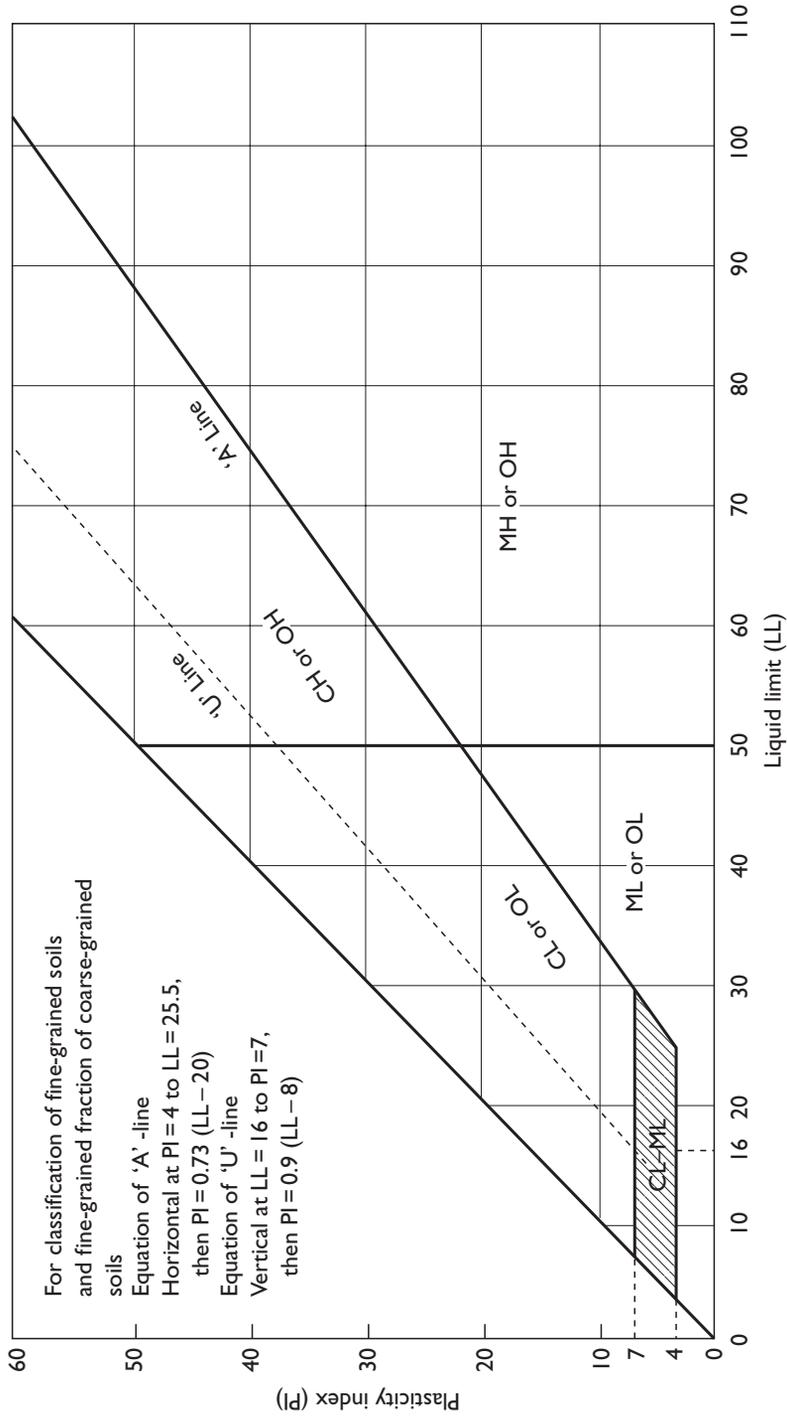


Figure 8.13 Unified classification system. (continued)
 Source: ASTM 2002.

Table 8.1 AASHTO soil classification system

General classification	Granular materials (35% or less passing 0.075 mm)			Silt-clay materials (more than 35% passing 0.075 mm)			
	A-1	A-3	A-2	A-4	A-5	A-6	A-7
Group classification	A-1-a	A-1-b	A-2-4	A-2-5	A-2-6	A-2-7	A-7-5, A-7-6
Sieve analysis, percent passing							
2.00 mm (No. 10)	50 max.	—	—	—	—	—	—
0.425 mm (No. 40)	30 max.	50 max.	—	—	—	—	—
0.075 mm (No. 200)	15 max.	25 max.	35 max.	35 max.	35 max.	35 max.	36 min. 36 min. 36 min. 36 min.
Characteristics of fraction passing 0.425 mm (No. 40)							
Liquid limit	—	—	40 max.	41 min.	40 max.	41 min.	40 max. 41 min. 41 min.
Plasticity index	6 max.	N.P.	10 max.	10 max.	11 min.	11 min.	10 max. 10 max. 11 min. 11 min. ^a
Usual types of significant constituent materials	Stone fragments, gravel and sand	Fine sand	Silty or clayey gravel and sand	Silty soils	Silty soils	Clayey soils	Clayey soils
General ratings as subgrade sand	Excellent to Good		Fair to poor				

Source: AASHTO (2002).

Note

^a Plasticity index of A-7-5 subgroup is equal to or less than LL minus 30. Plasticity index of A-7-6 subgroup is greater than LL minus 30.

References

- AASHTO (2002) *Standard Specifications for Transportation Materials and Testing*, Washington DC: American Association of State Road and Transportation Officials.
- Asphalt Institute (1989, reprinted 2000) 'The Asphalt Handbook', *Manual Series No. 4*. Kentucky: The Asphalt Institute.
- ASTM (2002) *Annual Book of ASTM Standards*. Philadelphia: American Society for Testing and Materials.
- BSI (1990). 'Methods of test for soils for civil engineering purposes', *BS 1377*. London: British Standards Institution.
- BSI (1984–95). 'Testing aggregates', *BS 812*. London: British Standards Institution.
- Lawrance, C.J., Byard, R.J. and Beaven, P.J. (1993) 'The terrain evaluation manual', *TRRL State of the Art Review 7*. London: H.M. Stationery Office.
- Lillesand, T. M. and Kiefer, R.W. (1979) *Remote Sensing and Image Interpretation*. London: John Wiley and Sons.
- Millard, R.S. (1993) 'Road buildings in the tropics', *TRRL State of the Art Review 8*. London: H.M. Stationery Office.

Tropical soils and rocks

Bent Thagesen

9.1 Introduction

Rock and soil

To the geologist the term ‘rock’ covers units of the earth’s crust formed by certain geological processes. The term ‘soil’ is used to describe materials produced by disintegration of rocks. Many civil engineers restrict the term ‘rock’ to materials that cannot be excavated without blasting. The term ‘soil’ is then used more widely to describe all other naturally occurring materials. The engineering definition may cause confusion since modern equipment has made it possible to excavate many soft rock types without using explosives. In this chapter, the geological terminology is applied.

Road construction

Most roads are constructed on soils or soft rocks. Cuttings are made through soils occurring in the road line. Embankments are constructed from soils taken from cuttings or nearby excavations. The road pavement is placed on the local soil or on fill materials brought in to the site. In mountainous regions, hard rock may occur in cuttings and rock fragments may be used as embankment fill. The pavement structure is constructed mainly from coarse-grained soils (sand and gravel). Hard rocks are usually quarried, crushed and graded to make aggregate for construction of heavy-duty road pavements. The stability of cuts and fills is dependent on the properties of the soils and rocks concerned. The structural design of the pavement is dependent on the bearing capacity of the subgrade and the strength of the paving materials. It is obvious that thorough knowledge of the properties of the rocks and the soils in and near the road line is the basis for appropriate road design.

Tropical soils

Some rock types, such as reef limestone, are more widespread in the tropics than in temperate regions. Deeply weathered rocks are common in wet tropical areas. The soil-forming processes in the tropics are more active and often continuous. As a result, the technical properties of tropical soils may be quite different from the properties that characterize soils from regions with a temperate climate. Many emerging and developing countries are situated in the tropics and this chapter describes briefly the principal rock types, the soil-forming processes, and the characteristic technical properties of the most widespread tropical soils. The aim is not to give a complete description of the rocks and soils in the tropics, but to highlight some of the geotechnical problems that may confront the road engineer working in these regions.

9.2 Rocks

9.2.1 Grouping of rocks

Rocks are divided into three groups: igneous, sedimentary and metamorphic (McLean and Gibble 1992). This classification indicates the mechanism through which the rocks were formed. Igneous rocks were formed by cooling and solidification of hot molten rock material (magma). Sedimentary rocks were formed by consolidation and cementation of sediments that have been accumulated in water or deposited by wind. Metamorphic rocks were formed by the modification of igneous or sedimentary rocks as a result of pressure, heat and also, occasionally, as a result of chemical action.

9.2.2 Igneous rocks

Igneous rocks may be divided into two groups: extrusive and intrusive.

Extrusive rocks are formed by rapid cooling of magma pouring out on the surface of the earth. The rapid cooling produces a fine-grained, often glass-like structure. Loose particles ejected from volcanic vents are called pyroclastic materials.

Intrusive rocks are formed when magma solidifies within the earth's crust. Cooling of large volumes at great depths takes place very slowly and results in rocks with coarse-grained, crystalline textures. Cooling of smaller masses of magma in cavities and cracks near the earth's surface takes place more rapidly and produces medium-grained rocks or porphyries where both large and small crystals are present. Intrusive rocks are often exposed as a result of erosion of overlying materials.

Igneous rocks are commonly classified according to silica content and texture, as in Table 9.1. Normally, coarse and medium-grained acid rocks are lighter coloured than their basic equivalents. Acid extrusive rocks, however, may be glassy and dark-coloured.

In general, igneous rocks make good road aggregates. Fine-grained types have better abrasion and impact values but poorer polished stone values than coarse-grained types of the same composition. Generally, the basic types have better affinity for bitumen than acid types. Dolerite, basalt and granite are the most common source of road aggregate among igneous rocks.

Extrusive
rocks

Intrusive rocks

Classification

Engineering
properties

Table 9.1 Classification of igneous rocks

Acidity	Amount SiO ₂ (%)	Type		
		Extrusive (fine-grained)	Intrusive	
			Medium-grained	Coarse-grained
Acid	>65	Rhyolite	Porphyries	Granite
Intermediate	55–65	Andesite	Porphyries	Diorite
Basic	<55	Basalt	Dolerite	Gabbro
	<65	Trachyte	Felsite	Syenite

Pyroclastics	Pyroclastic materials include dust, ash and cinders formed from lava blown apart by expanding gasses as the magma neared the surface. The dust and fine ash may be carried for great distances by the wind before the materials settle on land or in water. The coarse cinders travel a shorter distance. Consolidated dust and ash are called tuff, while consolidated cinders are termed agglomerate. Pyroclastic materials are normally porous and therefore have a low strength. These materials are not satisfactory for aggregate in concrete or asphalt, but experience from a number of countries indicates that various types of cinders and agglomerates make suitable base materials.
Weathering	Rocks disintegrate as a result of weathering. This includes physical disintegration caused by wind, water and freezing, and chemical disintegration caused by oxidation, reduction, soluviation and leaching. In the past, expansion and contraction because of temperature variations were thought to be an important agent of weathering, but it has now been proved that this process requires water to work (McLean and Gible 1992). The weathering processes are much more severe in the tropics than in regions with a temperate climate. High temperatures and – in wet tropical areas – high humidity make the processes more active and often continuous. Consequently, it is commonly found that rocks exposed on the surface in wet tropical areas are deeply weathered. The deterioration may not be obvious in a hand sample, but the consequences of using such weathered rock in a road pavement may be disastrous (Millard 1993). Therefore, igneous rocks in wet tropical regions should be subject to mineralogical investigation before a quarry site is opened, unless the quality of the rock is known from previous experience. It is not advisable to rely on the simple soundness test. The soundness test can be used to indicate whether rocks are weak and porous, but the test is less reliable for identifying rocks that are only partially weathered.

9.2.3 *Sedimentary rocks*

Sedimentary rocks are by far the most common rock type exposed at the earth's surface. The principal sedimentary rocks are sandstone, shale and limestone. Calcrete and silcrete are also mentioned.

Sandstone and shale	Sandstone and shale are rocks containing discrete particles derived from the erosion, transportation, and deposition of pre-existing rocks and soils. Sandstone primarily contains sand-sized particles, while shales are composed of clay-sized particles.
Limestone	Limestone is formed in water from the sedimentation of shells and shell fragments or chemically from the precipitation of calcium carbonate. Limestone that contains a significant amount of magnesium carbonate is called dolomite. Chalk is a porous limestone sometimes containing chert or flint nodules. Reef limestone is formed by corals, that is, organisms living in warm ocean water at moderate depth.
Engineering properties	A characteristic feature of many sedimentary rocks, particularly sandstone and shale, is a layered structure. The stratification is a result of variation in the depositional process. Hard sandstone and hard limestone are frequently used as aggregate in road pavements, but most sedimentary rocks are soft and only suitable for embankment construction. Hard sandstone has a good polished stone value but, in most cases, a substandard abrasion value. The affinity for bitumen varies. Hard limestone may have a high abrasion value and can then be used in sub-bases and bases.

Limestone bonds well to bitumen. However, most limestones have a very low polished stone value, and asphalt surfacings made with limestone aggregates normally become slippery when wet. Hard limestone is an excellent aggregate for use in cement concrete because of its low thermal expansion.

Limestone is widespread in many countries in the Middle East and may be the only material available to use as crushed rock aggregate. In other tropical regions, reef limestone is very common. In the South Pacific and Indian Oceans, many of the islands are formed entirely by corals. Several islands in the Caribbean and parts of the land mass surrounding the Bay of Mexico are formed by reef limestone raised from the seabed. Some of this limestone is hard and can be used for pavement construction (Millard 1993).

Sources of limestone

Calcrete and silcrete are chemical sediments produced by carbonate holding groundwater in particular localities in deserts and semi-deserts. The water is brought to the surface by capillary action where evaporation produces a hard precipitate. The composition of the deposits depends on the type of the underlying rock. Carbonates of calcium and magnesium are redeposited as calcrete, while silcrete is formed by siliceous precipitates. Massive calcrete resembles soft limestone while silcrete is similar to soft sandstone. Both calcrete and silcrete may also occur as nodular gravels. Calcrete and silcrete are normally only used for road making in areas where stronger materials are not available. Both materials may be stabilized with cement.

Calcrete and silcrete

9.2.4 Metamorphic rocks

The principal metamorphic rocks are quartzite, marble, slate, schist and gneiss. They are produced as a result of subjecting sedimentary rocks to very high temperatures and pressures over long timescales. Metamorphic rocks occur to only a limited extent.

Metamorphism of sandstone and limestone produces quartzite and marble respectively. Shale is altered to slate and schist, which both have a pronounced laminar structure (foliation). At the highest metamorphic grades, foliation becomes less distinct and the grain size coarser. This rock is called 'gneiss' and resembles coarse-grained granite.

Quartzite, marble, slate, schist, gneiss

Quartzite usually produces road aggregate with good abrasion and impact values. The polished stone value may be high, but the affinity for bitumen varies. Marble has properties similar to those of hard limestone; gneiss has properties similar to those of granite. The foliated metamorphic rocks, slate and schist, have very poor crushing strengths parallel to the banding and can only be used as a fill material.

Engineering properties

9.3 Soils

9.3.1 Soil types

There are three different soil-forming processes:

- Residual soils are formed in place by weathering of bedrock.
- Sediments are formed from parent materials that have been transported to their location by wind, water or glacial ice.
- Organic soils are formed from decomposed plant and animal materials.

The variation of tropical soils is extreme. Descriptions of the six most widespread groups of tropical soils follow. Organic soils create serious problems for road engineers in many temperate regions but are rare in the tropics, except in areas that are perpetually flooded. In hot wet climates, decomposed organic matters are quickly leached out of the soils. No organic soils are discussed.

9.3.2 Laterite

Laterite is a group of highly weathered soils formed by the concentration of hydrated oxides of iron and aluminium (Charman 1988). Iron oxides produce yellow, ochre, red or purple colours, but red is normally predominant. Other names for lateritic soils are roterde, ochrosols, ironstone (Nigeria), ferricrete (Southern Africa), mantle rock (Ghana) and murrum (East Africa). The term 'laterite' may correctly be applied to clays, sands and gravels in various combinations. However, a tendency to call all red tropical soils 'laterite' has caused much confusion.

Distribution

Laterite is a very widespread soil group. They occur in all wet tropical regions, including East, West and Central Africa, Brazil, Indonesia, Thailand and various islands, such as Hawaii and Cuba (see Figure 9.1).

Formation

Lateritic soils are residual soils formed in hot, wet tropical regions with an annual rainfall between 750 and 3,000 mm. The main soil-forming process consists of intensive weathering, with leaching of bases and silica, resulting in a relative accumulation of iron and aluminium oxides, and formation of kaolinitic clays. Intensive weathering produces deep laterite profiles that occur on flat slopes in the terrain where the run-off is limited. On level ground, where drainage is poor, expansive clays dominate at the expense of laterites. Two aspects of the parent rock affect the

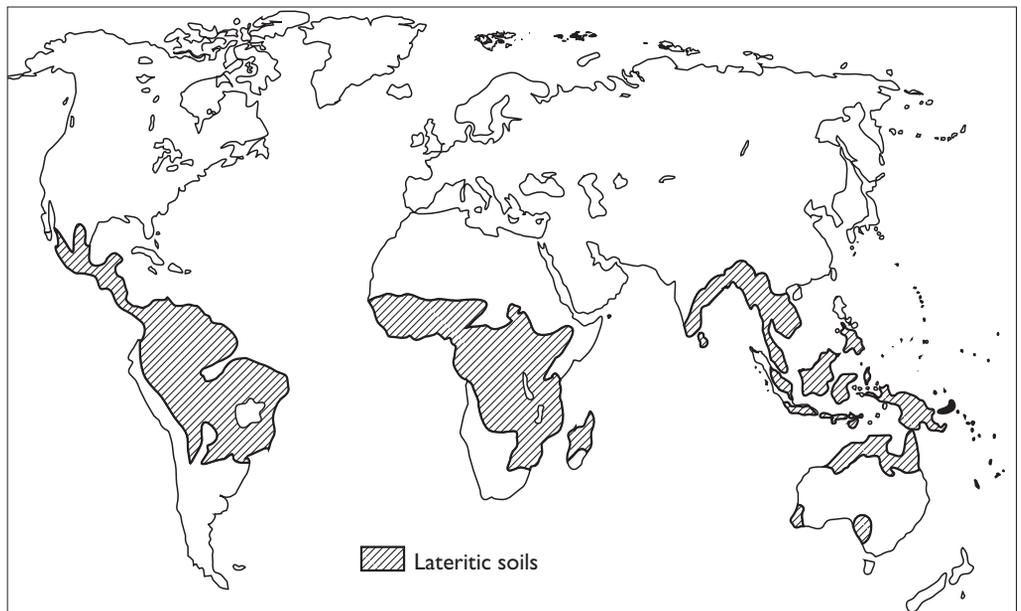


Figure 9.1 Distribution of lateritic soils.

Source: Charman (1988) (Reproduced by kind permission of CIRIA).

formation of laterite. One is the availability of iron and aluminium minerals. These are more readily available in basic rocks. The other is the quartz content of the parent rock. Where quartz is a substantial component of the original rock, it may remain in the weathered product as quartz grains.

If the iron-enriched laterite is dehydrated, hardening or concretionary development takes place. Dehydration may be caused by climatic changes or upheaval of land. Dehydration may also be induced by human activities, for example, clearing of forest.

Laterite mainly occurs as

- surface deposits of unhardened, clayey soils;
- massive rock-like hardpans;
- gravel consisting of concretionary nodules in a soil matrix.

Unhardened, clayey laterite is widespread in Southern Asia. It is this soil that has given name to laterite: meaning 'brick' in Latin. The first reference is from India, where this soft, moist soil was cut into blocks of brick size and then dried in the sun. The blocks became irreversibly hard by drying and were used as building bricks. Blocks cut from laterite were used to construct the famous deserted city of Angkor in Cambodia.

Rock-like crusts of hardened laterite occur in Southern Asia, Australia and Africa. Because of high resistance to erosion, these so-called hardpans are often found as cappings on detached plateau remnants (mesas), but their formation would have been in lower parts of slopes. The hardpans can be subjected to a new cycle of weathering and transportation in which hardpan fragments are re-deposited elsewhere as secondary deposits. Rock-like laterite is usually very heavy and abrasive. It is unsuitable for quarrying and processing for road aggregate.

Lateritic gravels occur in almost all tropical countries. They consist of gravel-sized concretionary nodules in a matrix of silt and clay. The resulting particle size distribution is gap-graded, as in Figure 9.2. Lateritic gravel deposits often stand out as low humps in the terrain. Typically, they take up an area of several hectares and a thickness between one and five metres.

The natural vegetation of lateritic soils in regions with high rainfall is dense, tropical rain forest. When the soils are cultivated, the organic matter is mineralized rapidly, and soil fertility becomes extremely low. In regions with less rainfall, lateritic soils mainly form bush savannah. Fertility is low because the hardened soil restricts growth.

In the tropics, where weathering is often intense, the availability of suitable rock as a source of crushed aggregate is often limited. In these areas, lateritic gravel is a traditional source of road aggregate. When the grading of lateritic gravel is close to a mechanically stable particle size distribution, the material performs satisfactorily on lightly trafficked roads, both as sub-base and base under thin asphalt surfacings, and as natural gravel surfacings. However, the significant silt and clay content often renders the material moisture-sensitive. The nodules are often hard with a good abrasion value but, in some deposits, the nodules are so weak that they can be crushed with the fingers. Many laterites contain a proportion of quartz. If the nodules are comparatively weak, the strength and durability of the coarser fraction become

Hardening

Occurrence

Unhardened
clay

Hardpans

Gravel

Vegetation

Engineering
properties

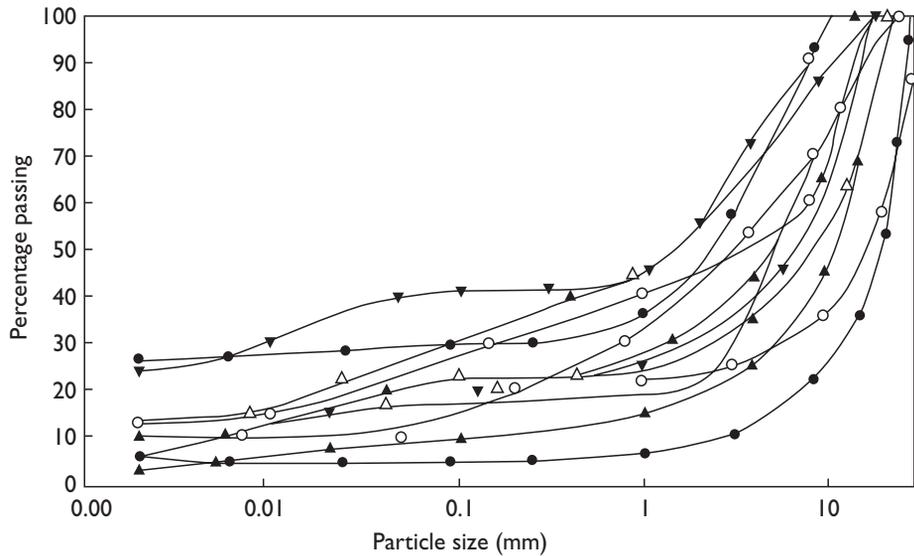


Figure 9.2 Typical gradings of lateritic gravels.

Source: Millard (1993).

a function of the proportion of quartz. The widespread evidence of the deterioration of laterite roads emphasizes the need for careful assessment of lateritic gravel that is to be used for road construction.

Self-hardening

According to Lyon Associates (1971), some deposits of lateritic gravel are immature and exhibit self-hardening properties when drying. This tendency may indicate that traditional requirements governing the selection of road materials can be relaxed. Thus, gravel that is normally considered mechanically unstable and too plastic may give acceptable performance as a pavement material. According to Millard (1993), the hardening of laterite pavements is actually due to a combination of causes. Compaction is improved by traffic and the construction process, since weak, coarse particles tend to fracture. This somewhat alleviates the gap-grading and improves the stability.

Stabilization

Laterite gravels tend to be gap-graded with a depleted sand-fraction, to contain a variable of fines, and to have coarse particles of variable strength, which may break down. These factors limit their usefulness as pavement materials on roads with heavy traffic. For heavy-duty pavements, performance needs to be improved by stabilization, which has proved successful with both cement and lime. In particular, satisfactory results have been obtained with cement.

Asphalt

Many cases have been reported of failure of new asphalt pavements placed on old laterite roads. This may be due both to the moisture sensitivity of the laterite, and to detrimental vapour pressures that may develop under the asphalt if the gravel base is very dense and humid, and the sun is shining on the black, heat-absorbing road surface.

Grading analyses and compaction tests

Particular care must be taken with the grading analyses of laterite. It is important that the sample preparation and test procedure do not fracture the particles. Furthermore, grading should be calculated by volume proportions as well as by mass proportion. This is because the coarse, iron-rich fraction in laterite usually has

a specific gravity of 3.0–3.5, while the fine, kaolinite fraction has a specific gravity of about 2.7. Breakdown of weak particles may affect the compaction characteristics, and fresh material should be used for each point on the compaction curve.

The los angeles abrasion test is commonly used to evaluate the particle strength. For lateritic gravels, however, the water absorption test may be used as a low-cost alternative. The water absorption is fairly consistently related to the abrasion value (see Figure 9.3).

Particle
strength

The plasticity limits of laterite can be misleading if it is not realized that a high proportion of hydrated oxides can change the properties of the fines. The oxides tend to coat the surface of the individual soil particles. The coating can reduce the ability of the clay mineral to absorb water. It can also cause a cementation of the adjacent grains. Both factors reduce plasticity, but intense remoulding of the soil breaks down the oxide coatings and the aggregations, with an attendant increase in the plasticity. When soils containing hydrated oxides of iron and aluminium are dried, they may become less plastic. Re-wetting cannot reverse the process. The effect is particularly pronounced when the material is oven-dried. The susceptibility of lateritic soils to the effect of clay mineral aggregation and to drying should be identified by determination of the plasticity limits after different mixing times, and by comparing the plasticity limits of soil prepared from natural moisture content with those of oven-dried samples re-wetted to the point of test (see Chapter 8).

Plasticity tests

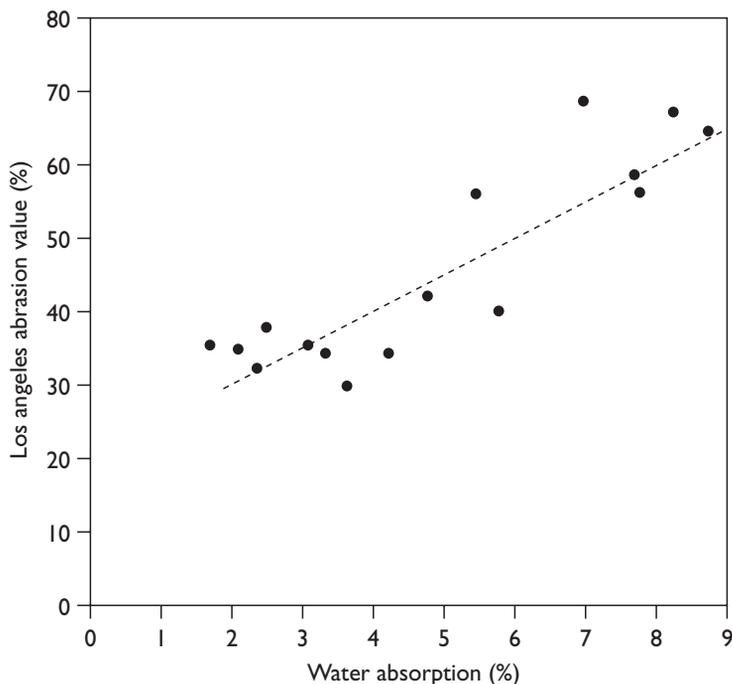


Figure 9.3 Relation between water absorption and los angeles abrasion value for typical West African lateritic gravels.

Source: Charman (1988) (Reproduced by kind permission of CIRIA).

9.3.3 Desert soils

Desert soils are normally called arid and semi-arid soils by geologists. Deserts are defined as regions with less than 100 mm average precipitation and semi-deserts as regions with less than 400 mm on an annual basis. However, the total precipitation is typically very irregular from year to year, between seasons of the year, and from place to place. Rain often falls in short heavy, local showers.

Distribution Desert soils occur mainly in the subtropics, although there is also some occurrence in dry regions of tropical countries. More than 30 per cent of the land surface of the earth consists of deserts and semi-deserts. They include the Kalahari, Namib, Sahara, Somalia, Arabia, Iran, Turkestan, Takla-Makan, Gobi, Patagonia, Atacama, the main part of Australia and part of North America. Dry regions of the world are expanding at a rate of 6 million hectares annually. The advance of deserts is considered as the biggest environmental problem on earth by many specialists.

Formation and characteristics The soil-forming process in dry regions is essentially the same as that in humid regions. However, because of the limited amount of water available in the soil profile, the reactions are relatively less intense. As a result, the soils inherit much of their composition from the parent material. The air is extremely dry in deserts and semi-deserts. When it is raining, much water is lost by evaporation and run-off. If there is no influence of groundwater, then the lower part of the soil is always dry. Especially in real desert soils, infiltration of rainwater is restricted because the wetting and rapid drying of the soil surface create a thin, hardly permeable crust on the surface. Therefore, most rainwater is lost as run-off causing sheet and gully erosion. The run-off finally enters deep wadis through which an enormous amount of water flows after a heavy shower. Common soil types in deserts and semi-deserts are sand and saline soils.

Sand Erosion and transport of soils by wind are common in desert regions, which lack a protective cover of vegetation and humidity to bind the soil grains together. If the original soil was gravelly, a layer of gravel, pebbles and stones is left behind on the surface forming a 'desert pavement' that protects the soil from further erosion. In other places, wind-blown material is accumulated on top of other soils. Sometimes it forms a thin sheet. In other places the wind-blown material forms drifts or dunes. Sand-dunes differ from sand-drifts in that no fixed obstacles are necessary to initiate the formation of sand-dunes. About 15 per cent of the surface of the Sahara is covered by drifting sand-dunes. Other large bodies of drifting sand are found in Saudi Arabia. The material forming desert drifts and dunes is normally fine quartz sand.

Saline soils In low-lying areas, particularly on the coast, saline groundwater may be brought to the surface by capillary action. Evaporation causes the salt to accumulate in the upper layer of the soil. The types, mixture and concentration of salt vary. Sodium chloride is often less than half the total, which may also include sodium sulphate and chlorides of calcium and magnesium. Some saline soils also contain carbonates. Nitrates and borates are more rare. The salts may occur as salt crystals, salt crusts and salic horizons. The soils have mostly a loose porous granular structure. If hygroscopic salts such as calcium and magnesium chlorides are present, they attract moisture from the air making the soil look moist in the morning. If salt reaches the surface a salt crust may form. Saline soils are common in some deserts, particularly in the Sahara and Saudi Arabia.

<p>The natural vegetation on non-saline desert soils is scanty grass growth and small bushes. Overgrazing and cultivation of soils that are too dry have caused severe wind erosion in many desert regions. Saline soils support very little vegetation. Irrigation of arid and semi-arid soils involves a risk of salinization so that the land is damaged and made unproductive.</p>	Vegetation
<p>Desert sands usually occur at low field densities. When constructing roads on loose sand, it is necessary to compact the sand thoroughly to avoid uneven settlement of the road. However, it may be difficult to obtain satisfactory compaction, because water is scarce and dry compaction is not very effective with the often single-sized soil material. Vibratory compaction can be helpful in obtaining in-depth compaction, and this can be followed by static compaction to settle the top layer.</p>	Engineering properties
<p>When locating roads in desert regions, a problem is to avoid drifting sand. In principle, this can be controlled in the same way as drifting snow in regions with cold climate. Road cuttings in blown sand should be avoided since they act as sand traps. The road should be placed on the windward side of obstacles so that sand is blown clear of the road. The cross-section of the road should be raised slightly and made as smooth as possible. Side drains should be avoided if possible. In some places, establishment of vegetation has been used to combat sand drift.</p>	Drift control
<p>The single-sized particle distribution make desert sands unsuitable for stabilizing with cement, but they can often be treated effectively with bitumen.</p>	Stabilizing
<p>Salt-holding sand can be used as a fill and base material if special precautions are taken. Hygroscopic salts in soils have a stabilizing effect. If the salt content is high, however, the stability is drastically reduced when the soil takes up water. Therefore, embankments and bases made from saline materials should normally be protected from being soaked during rain.</p>	Use of saline soils
<p>Salts have no destructive chemical effect on asphalt, but a high salt concentration in the pavement, subgrade and/or groundwater may result in blistering of thin asphalt pavements. The blistering is caused by salts migrating upwards and crystallizing below the asphalt. The salt is precipitated from evaporating capillary moisture. Thicker asphalt pavements (30mm or more) do not appear to be affected by salt. They are more resistant to high evaporation rates and crystal pressures (Obika and Freer-Hewish 1990).</p>	Asphalt
<p>Sand and gravel contaminated with salts should normally be avoided as aggregate in concrete. Sulphates are particularly detrimental. In the Middle East, several concrete structures have disintegrated because of sulphates in the aggregates. Culverts, bridge foundations and other concrete structures, covered by saline soil, should be made of concrete with low permeability and protected from seeping ground water.</p>	Concrete
<p>There are no special difficulties in performing classification and engineering tests on desert soils. Salt contents may be determined by use of traditional titration tests.</p>	Testing

9.3.4 *Expansive clays*

The most well-known example of expansive clay is 'black cotton' soil. This name is believed to have originated from India where locations of these soils are favourable for growing cotton. Many other names are applied locally, such as 'margalitic' soils in Indonesia, 'black turfs' in Africa and 'tirs' in Morocco. Expansive clays swell when moistened and shrink when dried. The swelling properties are due to a high

content of the clay mineral montmorillonite. The colour is black, dark grey or dark grey-brown. The dark colour is a result of a small amount of organic matter being mixed with the clay.

Distribution Expansive clays occur in a great many tropical and subtropical countries. The most extensive regions with expansive clays are found in India, Australia and Sudan. Less extensive areas occur in almost all tropical and subtropical countries, especially in Africa (see Figure 9.4).

Formation The parent material of most expansive clays is transported material, although some types are formed in situ, from weathered basalt or limestone. The climate needs to be warm, with alternating dry and wet seasons, and an annual rainfall ranging from 300 to 1,000 mm. Most expansive clays form in poorly drained areas.

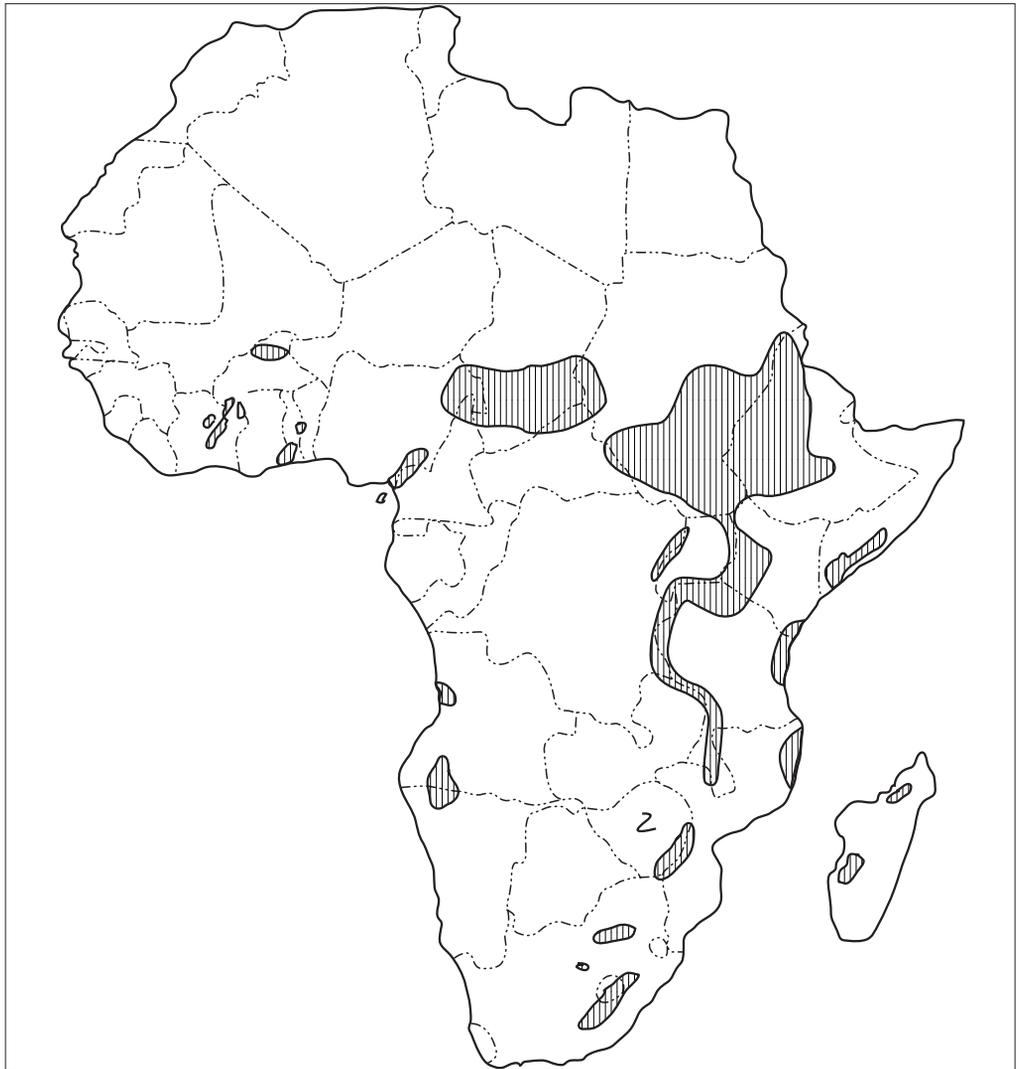


Figure 9.4 Distribution of expansive clays in Africa.

Source: Lyon Associates (1971).

Expansive clays rarely develop thick profiles, with the layer thickness normally varying between one and two metres (Buringh 1968).

During the dry season cracks develop in the surface of expansive clays as a result of shrinkage. The cracks may be 10 mm wide and more than 500–800 mm deep. The cracks open up in a polygonal pattern, each polygon having a diameter of 1–4 m. The higher the clay content, the smaller the polygons. During the dry season, small particles from the surface fall into the cracks. During the following wet season, when the soil swells, the partly filled cracks cannot close and subsoil is pressed to the surface. In this way, small mounds are formed, and former cracks become small depressions. Many expansive clays have a loose granular surface due to the cracking and swelling in the upper few centimetres of the surface. This process is called ‘self-munching’.

Characteristics

The natural vegetation on expansive clays is bush or grass savannah. In many countries, expansive clays are important agricultural soils because of their high fertility. However, the soils are difficult to cultivate because they are very hard in the dry season and very sticky in the wet season.

Vegetation

The California bearing ratio (CBR) of swelling soils is very dependent on the water content. This may be illustrated by an example from Africa, where an asphalt road was constructed on an expansive clay subgrade. After three months of dry season, the registered water content was 20 per cent in the subgrade underneath the asphalt surfacing, and 8 per cent below the unsealed shoulders. The corresponding CBR values were 30 and 3 per cent, respectively.

Bearing capacity

Swelling of expansive clay depends on the content of the clay mineral ‘montmorillonite’, the initial moisture content, and the amount of water that the clay is allowed or able to absorb. It is difficult to evaluate the potential swelling by simple laboratory testing, as the results are very sensitive to the test conditions. If an expansive clay is allowed to absorb water and at the same time prevented from expansion, a high swelling pressure develops. If a small expansion is allowed, the expansion pressure is reduced considerably.

Swelling

The liquid limit of most expansive clays is over 50 per cent. The plasticity index normally falls between 20 and 60 per cent. The liquid limit related to the clay content may be used as an indicator of the swelling potential (see Figure 9.5). A high liquid limit, however, is not a proof of expansiveness. Also, non-expansive clays may have a high liquid limit.

Plasticity

The free swell test is a simple method to verify swelling tendencies (see Chapter 8). It provides no quantitative measure of either pressure or volume change but, nevertheless, is useful in identifying problem soils. The test results on a number of expansive clays from Africa varied between 50 and 100 per cent.

Free swell

Expansive clays with a high swelling potential are a persistent problem in road construction in tropical regions with pronounced dry and wet seasons. Road construction alters the moisture pattern in the subgrade soil. Surface evaporation is reduced by the road pavement and, after the road has been completed, the moisture content of the subgrade normally rises. This causes swelling of the subgrade and heaving of the pavement. The swelling also reduces the bearing capacity of the soil. The heaving of the pavement may be considerable if the pavement is laid when the subgrade soil is in a dry state. During the following wet and dry seasons, the surface of the road will move up and down depending on the moisture changes. Movements

Engineering properties

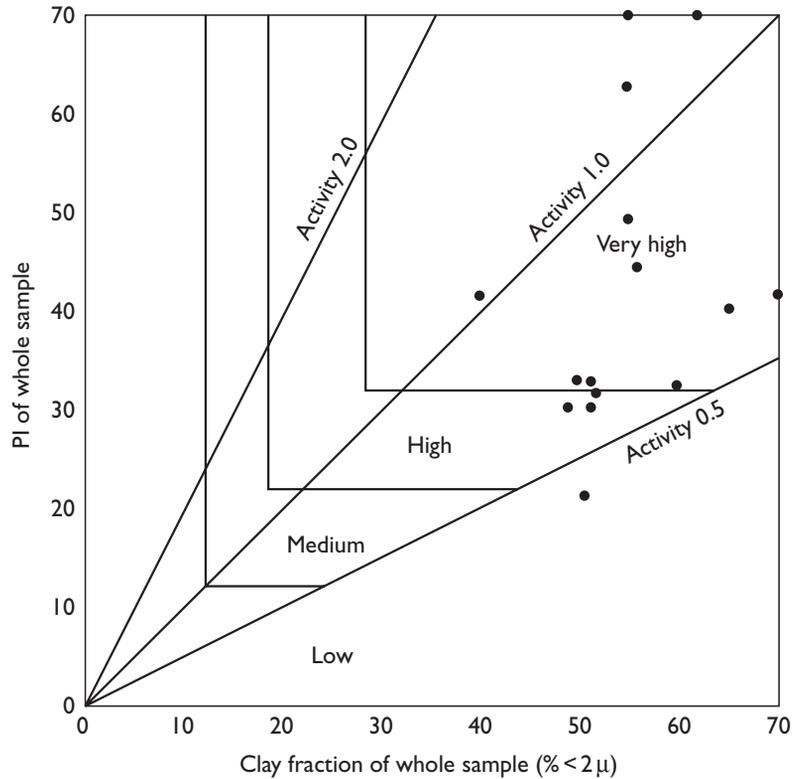


Figure 9.5 Potential swelling tendencies of clays.

Source: Lyon Associates (1971).

will be less pronounced at the centre of the road than near the shoulders and the side drains where water is able to infiltrate and evaporate. In highly expansive clays, the yearly vertical movements of the pavement edges may be as high as 50–100 mm, causing severe edge failures.

Utilization There are no inexpensive means of utilizing highly expansive clays satisfactorily. The obvious solution is to avoid them wherever possible or, if shallow, to excavate and dispose of the material. However, in many areas where these soils occur, the deposits cover such large areas that avoiding or bypassing them is not feasible, and there are no materials suitable for fill.

Base thickness Roads over expansive clays require a considerable thickness of base and sub-base, because of the low CBR values, even in the absence of swelling.

Drainage Drainage is extremely important. Poor drainage results in big seasonal variations in the subgrade moisture. However, if ditches are too deep, or too close to the pavement structure, seasonal drying or partial desiccation may be pronounced along the shoulders. The best solution is to locate the ditches at some distance from the pavement and seal the shoulders.

Pre-wetting Pre-wetting of the subgrade before placing the sub-base may reduce the heave in new pavements. Pre-wetting has been used with varying success.

The seasonal expansion and contraction of the subgrade may be minimized by confining the subgrade under an embankment. However, this approach assumes that suitable embankment material is readily available. **Embankment**

Expansive clays are normally very suitable for lime stabilization. Lime stabilization increases bearing capacity and reduces the potential swelling in the top layer of the subgrade. The treated layer is more impermeable, so moisture variations are reduced. Several cases are reported where admixture of 4–6 per cent of hydrated lime has made an expansive clay suitable as support for a traditional road pavement. **Lime stabilization**

9.3.5 Volcanic ash soils

Volcanic ash soils are residual soils formed in tropical regions with current or recent volcanic activity. Volcanic ash soils occupy large areas in Central America, Western South America, Hawaii, New Zealand, Japan, the Philippines, Indonesia, some East African countries and Cameroon. In Indonesia, these soils are known as ‘asgronden’ or ‘zwarte stofgronden’. **Distribution**

When volcanic ash weathers, it retains much of its original loose structure. In hot humid climates, the weathering forms highly plastic clay containing allophane, an amorphous mineral. In regions with constant high humidity, the allophane stage can persist for long periods. In climates with distinct dry seasons, the allophane stage is followed by development of the clay mineral halloysite (Millard 1993). **Formation**

Allophanic clays have a very fragile structure that may collapse when loaded. Whereas the particle shape of other clay minerals is flat and plate-like, the particles in halloysite are curled and twisted. The presence of allophane and the porosity of the pumice-like parent material cause a high water-holding capacity. The natural water content may be more than 200 per cent. If allophanic clay is heated, the small clay-sized particles develop aggregations, and the plastic clay changes into a non-plastic, sandy soil. In halloysitic soils the curled clay particles help create an open structure. The structure may be reinforced by precipitated iron oxides. A typical example is the red-coffee soil occurring in East Africa. These soils are free-draining, and the structure is much stronger than in allophanic clays. Heavy working, however, may break down their structure irreversibly. **Characteristics**

Volcanic ash soils have a high natural fertility and they are among the best agricultural soils in the tropics. They support different types of crops. **Vegetation**

Allophanic clays are highly sensitive to disturbance, so road engineers should avoid working in their vicinity. Roads can be constructed over halloysitic clays, but heavy compaction of the soil should be avoided. This can break down the soil structure and render the soil weaker and more susceptible to the effect of water. Studies (Nicholson *et al.* 1994) have demonstrated the potential for lime modification of halloysitic clays. **Engineering properties**

Sieve analysis, plasticity tests and compaction tests are likely to break down the structure in halloysite with the effect that the test results do not represent the engineering properties of these soils. For the same reason, CBR tests on remoulded samples may not indicate the true bearing capacity. The compaction properties and the design CBR of halloysitic clay soils should rather be established from field trials or local experience. **Testing**

9.3.6 Tropical alluvial soils

Alluvial soils are mineral soils that have been transported and deposited by flowing water.

Distribution

Alluvial soils are scattered all over the world. They are not specific to the tropics, but deserve mention because the coarser types are important sources of road making materials in many countries. Big rivers and deltas in the tropics with vast alluvial deposits are the Indus, Ganges, Brahmaputra, Mekong, Irrawaddy, Digul, Congo, Niger, Amazon and Orinoco. Coastal flats with clay sediments are well-known in Surinam, Guyana, Sierra Leone and South-East Asia. Irrigational sediments, now forming alluvial soils, are best known in Iraq.

Formation

Alluvial soils mainly occur in

- river plains and deltas;
- alluvial fans;
- former lake bottoms;
- old irrigated areas;
- coastal plains.

River plains and deltas

River plains and deltas are landforms created by the processes of stream-meandering and over-bank flooding. Most alluvial soils form flat areas in the lowest parts of the landscape, and they are normally wet with a high water table, permanently or during specific periods. River banks are eroded on the outside and downstream side of each bend. On the inside of each bend, sand and gravel are usually deposited. These deposits are called 'point bar deposits'. In the course of time, a bend in the river bulges more and more, and a point is reached where the river changes its course. The bend is cut off, and a natural basin is formed. When the river is overflowing, its bank material is deposited outside the normal stream channel as natural levees. In deltas, the pattern of deposition is similar to that of flood plains, but the rivers have more branches.

Alluvial fans

Alluvial fans are deposited at the base of steep slopes where streams discharge into an area having a more gentle relief. The fans are made up of coarse-textured materials eroded from the higher elevations and transported downslope. Alluvial fans are most widely developed in arid areas at the base of high mountain ranges where they may be several kilometres long. In humid areas, fans are smaller, typically only a few hundred metres long.

Lake bottoms

The lacustrine pattern of sedimentation is less distinct. Usually, soils are uniform and clayey, often mixed with organic material.

Coastal plains

Coastal plains are deposited in seawater, and influenced by the tides. The saline clayey soils are often called 'marsh soils'. Near the coasts, there are former beach walls comprising sand, gravel or shells.

Characteristics

Alluvial deposits are normally stratified with a uniform particle size in the individual stratum. The particle size depends on the velocity of water during deposition. The mineral composition is heterogeneous, reflecting the geology of the parent materials upstream in the catchment area. They may consist of a mixture of various unweathered, weathered and partly weathered minerals depending on the cycles of erosion and re-deposition that have taken place. In wet and humid tropical regions,

the sediments often consist of almost completely weathered material. In volcanic regions, some fresh minerals may be present. In arid and semi-arid tropical regions, the sediments are normally made up of less weathered minerals, often high in carbonates and gypsum. Marine sediments usually contain weatherable minerals. Fine-textured clay sediments deposited under water contain much more water than normal soils, and the packing of the particles is very loose. When water drains off or evaporates, there is a gradual vertical shrinkage. On drying, air penetrates the soil through cracks, and the bluish-grey colour changes to brownish as a result of oxidation. Other chemical and biological processes also take place. The transformation of the clay is irreversible. It will never take up all of the lost water even if submerged again.

The weathered alluvial soils from the wet and humid tropics are not usually as fertile as the less weathered alluvial soils in temperate regions. The most common crop is rice, but sugar cane, bananas and, on better drained sites, cocoa, coffee and citrus are also grown.

Vegetation

Good quality road construction materials and aggregates can usually be extracted from point bar deposits in flood plains and from alluvial fans. When surveying gravel and stone deposits, it should be realized that, if stratified materials with different grain size in the individual layers are mixed, they may take up a volume 25 per cent less than the combined volumes of the layers. The smaller grains will partly fit into the space between the coarser grains. It is often necessary to process the materials to satisfy the requirements for aggregate to be used in granular bases, cement or asphalt concrete bases, or asphalt concrete surfacing. The processing may include washing, sieving and/or crushing.

Engineering properties

Traditional testing methods are used to investigate the engineering properties of alluvial soils. However, it should be remembered that the properties of fine-textured clay sediments may change irreversibly when dried and oxidized.

Testing

References

- Buringh, P. (1968) *Introduction of the Study of Soils in Tropical and Subtropical Regions*, Wageningen: Centre for Agricultural Publishing and Documentation.
- Charman, J.H. (1988) 'Laterite in road pavements', *Special Publication 47*, London: Construction Industry Research and Information Association.
- Lyon Associates (1971) *Laterite and Lateritic Soils and Other Problem Soils of Africa: An Engineering Study for the US Agency for International Development*, Kumasi: Building and Road Research Institute.
- McLean, A.C. and Gibble, C.D. (1992) *Geology for Civil Engineers*, London: Chapman and Hall.
- Millard, R.S. (1993) 'Road building in the tropics', *State of the Art Review No. 9*, London: H.M. Stationery Office.
- Nicholson, P.G., Kashyap, V. and Fujii, C.F. (1994) 'Fly ash admixture improvement of tropical Hawaiian soils', *Transportation Research Record No. 1440*, Washington, DC: Transportation Research Board.
- Obika, B. and Freer-Hewish, R.J. (1990) 'Soluble salt damage to thin bituminous surfacings of roads and runways', *Journal of the Australian Road Research Board No. 4*.

Hydrology and drainage

*Tue Hald, Jan Hassing, Michael Høgedal
and Alan Jacobsen*

10.1 Introduction

Water impact
on roads

One of the most important aspects of the design of a road is the provision made for protecting the road from surface water and groundwater. Water on the pavement slows traffic and contributes to accidents from hydroplaning and loss of visibility from splash and spray. If water is allowed to enter the structure of the road, the strength and deformation resistance of the pavement and subgrade will be weakened, and it will be much more susceptible to damage by the traffic. Water can enter the road as a result of rain penetrating the surface, or as a result of the infiltration of groundwater. When roads fail, it is often due to inadequate drainage. Water can also have a harmful effect on shoulders, slopes, ditches, and other features. Failures can arise spectacularly as, for example, when cuttings collapse or when embankments and bridges are carried away by flood-water. High water velocities can cause severe erosion, possibly leading to the road being cut. On the other hand, low velocities at drainage structures can lead to silt being deposited which, in turn, can lead to a blockage. Blockage often results in further erosion or overtopping and possibly wash-out.

Functions of
drainage

The drainage system has four main functions:

- to convey stormwater from the surface of the carriageway to outfalls;
- to control the level of the water-table in the subgrade beneath the carriageway;
- to intercept ground and surface water flowing towards the road;
- to convey water across the alignment of the road in a controlled fashion.

The first three functions are performed by longitudinal drainage, in particular side drains, while the fourth function requires cross-drainage structures, such as culverts, fords, drifts, and bridges.

Basis of design

The hydraulic design of longitudinal drainage is normally based on discharges calculated from local rainfall in the vicinity of the road and from further afield. The design of drainage structures involves choices about acceptable maintenance cost and risks versus costs of constructing drainage and erosion protection. Drainage often represent a significant proportion of the total road construction costs. However, drainage has a substantial impact on lifetime road performance.

Hydraulic
analysis

The first step in the drainage design procedure is the hydrological analysis. This analysis provides design discharges for all major drainage structures, and for rivers

and streams adjacent to the road alignment. Hence, the hydrological analysis involves characterizing the project area, for example, assessing the impact of rivers, water-courses and wetlands, assessing soil properties, vegetation zones and land-use in the catchment area and assessing the impact of rainfall. Several different types of hydrological investigations can lead to estimates of the design discharges needed. The type of investigation made will depend on the data availability and the possibilities for acquisition of supplementary data. The most comprehensive hydrological analysis is carried out using numerical rainfall run-off models such as Mike 11, which has been under continuous development since 1972. The Mike 11 model simulates surface flow, inter-flow and groundwater base flow, by taking into account time-varying storage capacity, irrigation and groundwater pumping. Numerical simulations have proved to be a reliable design tool.

The following problems often occur with hydrological design:

Problems

- rain gauge and discharge station networks are of low density, and the quality and accessibility of data are poor;
- few rainfall intensity measurements are available;
- only short time-series data exist;
- flood estimation practices are not standardized;
- land-use changes, high intensity rainfall and erosion cause dynamic changes in the drainage systems, making predictions difficult;
- cost constraints, non-availability of detailed maps and other basic data, and difficult logistics limit the quality and extent of hydrological surveys.

10.2 Rainfall

Although some regional and global patterns of rainfall can be identified, rainfall is most often unpredictable and highly variable both in time and space. The rainfall volume in a storm is the basic input to many flood estimation methods. The requirements of the method may vary from knowledge of the average annual rainfall to knowledge of the detailed variation of rainfall intensity during a rainstorm with a selected recurrence interval. The variability in volume and intensity during several rainfall events is depicted in Figure 10.1. Large variations in monthly maxima and minima occur and even, in some cases, the daily maximum constitutes a major proportion of the monthly maximum.

Variability

The extreme events that cause floods are often rainstorms with a duration of between half an hour to a few hours depending on the catchment size. For the larger catchments which have long response times, rainstorms of longer duration with less intensity will be most critical; for small catchment areas with short response time, short-duration high-intensity rainstorms will be most critical. Hence, knowledge of rain duration is also important.

Storms

Rainfall is recorded using daily rainfall or autographic gauges. An autographic gauge records both time and accumulated rainfall either graphically or digitally. The accuracy of measurements is influenced by the wind and any obstructions in the vicinity of the gauge. The wind effect will underestimate the actual rainfall. The underestimation can be 10–15 per cent depending on wind speed and exposure.

Rainfall recording

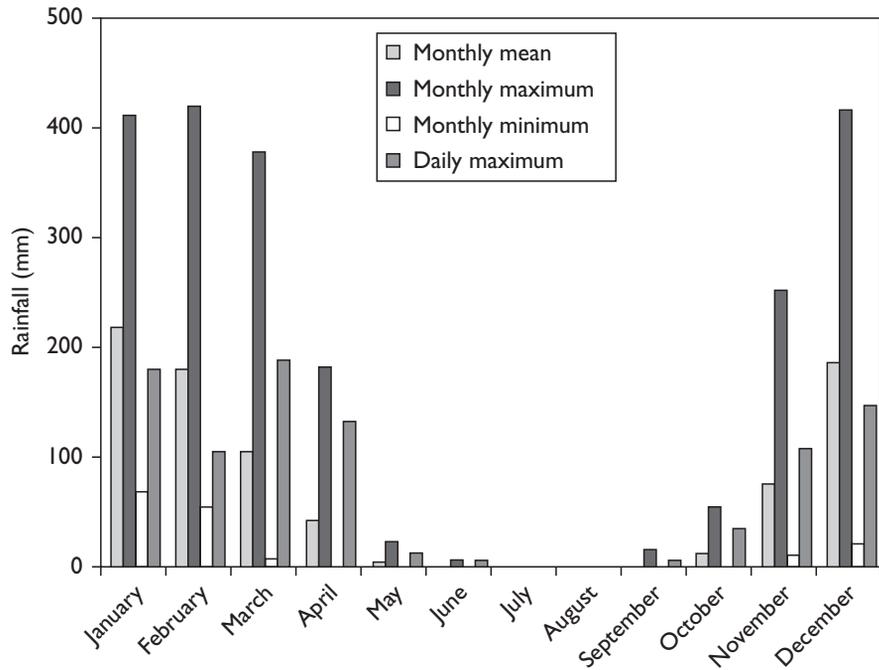


Figure 10.1 Monthly rainfall distribution in Lusaka, Zambia, 1974–2000.

Copyright: Rambøll.

Point and area rainfall

Measurements made by the above methods will represent the point-rainfall: the rainfall at a particular location. However, most often the total rain falling on a particular area is needed to assess discharges. The conversion of point-rainfall to area-rainfall for a selected standard duration (year, season, month, day, or fraction of a day) can be done by various methods. If data from several rain gauges are available in the catchment under study, then the following methods can be applied:

- *Averaging* – a simple average of the amounts of rain at the different gauges for the period of interest.
- *Theissen method* – a weighted average; neighbouring stations are connected on a map with lines, and perpendiculars are raised in the middle of each connecting line; each station is represented by an area bounded by the perpendiculars.
- *Isohyets* (curves representing the same amount of rainfall) – areas between isohyet curves can be measured, and the total rainfall calculated by multiplying the areas by the mean value of the higher and the lower isohyets.

The isohyet method is particularly useful in mountain areas, where they can be drawn to take into account the effects of topography.

Area reduction factors

High-intensity short-duration storms normally only affect areas of a limited size. For areas larger than about 10 km², this has to be taken into account in rainfall runoff modelling by applying reduction factors to the point-rainfall. In general, the area reduction factor (the conversion between point-rainfall and area-rainfall) is

dependent on both the area and the duration, as well as on the ‘severity’ of the storm. Severity is expressed as a recurrence interval between rain events. However, dependency on the recurrence interval is often not significant. An example of an area reduction factor (ARF) is ‘TRRL East African flood model’ validated for Nairobi (Watkins and Fiddes 1984):

$$\text{ARF} = 1 - 0.2 \times t^{-1/3} \times A^{1/2}$$

where t is the rainstorm duration in hours; and A the area in km^2 . This equation applies for storms of up to 8 h duration. No influence of storm severity is applied. An example of a more general average relationship for reducing the area is shown in Figure 10.2 for a rainstorm with a 5-year recurrence interval.

The high variability associated with the rainfall requires a probabilistic approach to describing rainfall characteristics. The worst rainstorm occurring in, say, 50 years has a 50-year interval or 50-year ‘return period’. The ‘recurrence interval’ (T) is defined as the average number of years between a rainstorm of a given size, and the recurrence interval is the inverted probability of occurrence. It should not be inferred that the storm occurs regularly at T -year intervals, but that it has a probability of occurrence of $1/T$ in any year. It can be shown that the probability of the occurrence of the T -year event within a period of T years is 63 per cent.

When analysing a set of rainfall data, the most widely applied tool is an ‘extreme value statistical analysis’. This method is also applicable to other data series, such as river discharge data.

The analysis is normally performed for an annual series of data. For a selected duration of rainfall, for instance the 24-h, 2-h, 1-h or 30-min rainfall, the highest value is determined from each calendar year. This series is termed an ‘annual

Recurrence
interval

Extreme value
analysis

Annual
maximum
series

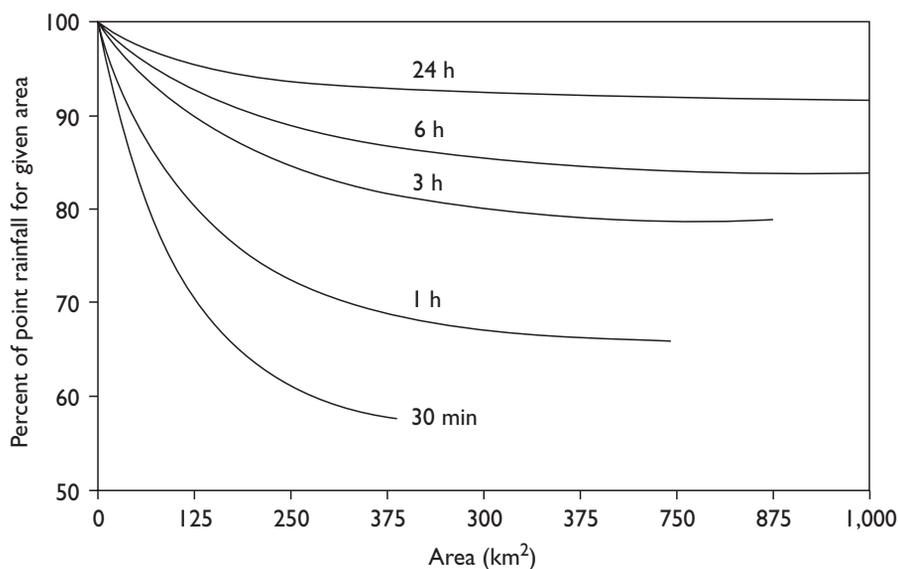


Figure 10.2 Area reduction for a 5-year recurrence interval.

Source: WMO (1970).

maximum series'. This approach can result in the selection of two rainfalls within the same wet season (e.g. one in December and one in January), and sometimes none from another wet season. This problem occurs when the wet season extends over the new-year, for example from October to April. This can be overcome by taking the 'hydrological year' instead of the calendar year, so that one peak is selected for each wet season.

Analysis method

The annual records are listed in order of descending values and a rank is assigned to each of them, the first having rank '1', the second rank '2', and so on. Next, the 'non-exceedence probability', or recurrence interval/return period, for each value is calculated by the weibull formula:

$$F = 1 - \frac{1}{T} = 1 - \frac{n}{N + 1}$$

where F is the non-exceedence probability; T the the recurrence interval; N the number of years in the series; and n the rank.

A distribution curve is fitted to the series of annual rainfall extremes. The best fit is usually obtained by using the 'gumbel distribution', but a log-normal distribution may in some cases be found applicable:

Gumbel

$$F = \exp\left(-\exp\left(-\frac{r-B}{A}\right)\right)$$

Log-normal

$$F = \Phi\left(\frac{\ln r-B}{A}\right)$$

where F is the non-exceedence probability or cumulative frequency of rainfall, for example, rain depth (mm) or intensity (mm/h); r the rainfall depth (mm) or intensity (mm/h); A, B are distribution parameters; and Φ the standard normal distribution function.

The distributions can be fitted using standard methods, such as 'maximum-likelihood' or 'least-squares'. Once the relationship is established, extreme values for other, normally longer, recurrence intervals can be calculated by entering the appropriate non-exceedence probability.

Station-year technique

A meaningful statistical analysis requires at least ten years of reliable records and, even then, extrapolated values need to be used. If records are shorter than 10 years, the 'station-year' technique may be applied. This technique mixes records from several stations, assuming that the records are homogeneous and that the events are independent. Records from several stations are then lumped together, considered as a single record and analysed accordingly. The results are then considered valid for all the stations that were sources of data. The station-year technique may be appropriate for rainstorm analyses, but should not be used in an analysis of discharge data. An example of an extreme-value probability distribution is shown in Figure 10.3.

Intensity-duration relations

It is unusual for adequate rainfall data to exist for the development of a full set of rainfall intensity curves for the location of interest. Normally, only rainfall data for a fixed duration, such as the 24-h rainfall, is available and no information on the time

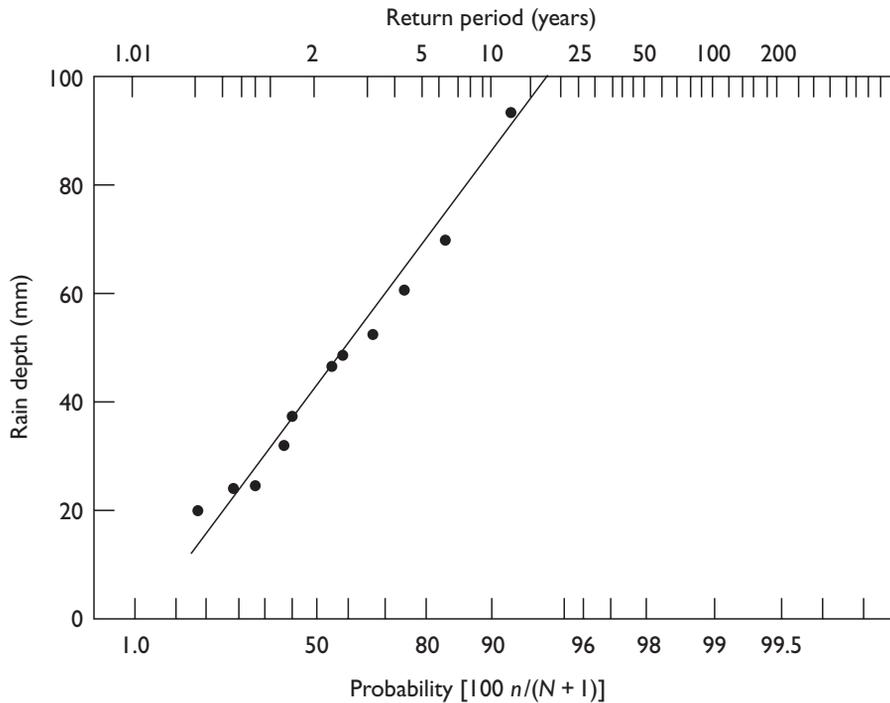


Figure 10.3 Extreme-value probability distribution function.

Source: WMO (1970).

distribution exists. In such cases, it may be possible to use a generalized relationship and determine the appropriate constants to use in this. For a large number of rainfall gauges in East Africa, for example, the rainfall has been related to duration by applying the relationship:

$$RR_t = \frac{t}{24} \left\{ \frac{(b + 24)^n}{(b + t)^n} \right\}$$

where RR_t is the rainfall ratio $R_t : R_{24}$; R_t the rainfall in a given duration t (mm); R_{24} the 24 hours rainfall (mm); t the duration (h); and b , n are constants. The average value of b is 0.3 and n is 0.9 (range of n is 0.78–1.09).

The rainfall is easily converted to intensity by dividing by the corresponding duration. Examples of intensity-duration curves for different recurrence intervals are shown in Figure 10.4.

10.3 Flood discharge estimation

10.3.1 Background

The run-off in rivers and streams is the result of several hydrological processes. In general, run-off comprises the surface run-off, the inter-flow and the groundwater flow. The surface run-off is the direct result of the excess rainfall, while inter-flow is the water moving through the soil near the surface. The groundwater component is

Run-off
components

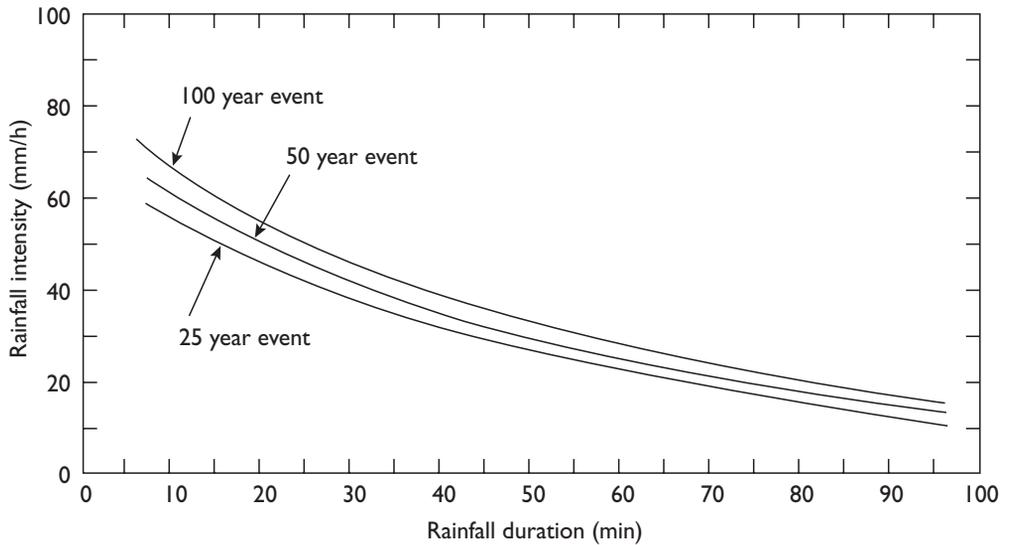


Figure 10.4 Intensity–duration curves.

the contribution from the aquifers through which the stream passes. During a major flood, the surface run-off is predominant and usually only the contribution from the excess rainfall has to be taken into account. The variation in discharge over time, as a result of the run-off processes, depends on a number of characteristics of the rainfall, the catchment and the stream. The most significant parameters are the rainfall intensities and their geographical distribution, the catchment area size, the shape and slope of the catchment, the soil infiltration capacity, the vegetation, the slope and roughness of the stream, and the storage capacity of the catchment.

Hydrograph

A plot of the discharge variation over time is called a ‘hydrograph’, and the peak of this hydrograph is the maximum discharge. For small catchments, maximum discharge occurs when rain falling on the remotest part of the catchment reaches the stream or drainage system. At this point in time, called the ‘time of concentration’, the whole catchment contributes to the discharge. For larger catchments, the run-off will be delayed before reaching the aquifer. The shape of the hydrograph varies significantly depending on the rainfall intensities, durations and on the catchment characteristics. The most peaked hydrographs are found in mountain catchments, where no vegetation attenuates the surface run-off, and where the infiltration capacity is very small. Lower peaks are found in low-relief catchments with extensive vegetation and good infiltration capacity. These catchments serve as reservoirs and will delay the run-off significantly. A hydrograph for a combination of mountain catchments (total area $A = 140 \text{ km}^2$) is shown in Figure 10.5.

Flood discharge estimation

Different methods for estimation of flood discharges have been developed. The methods vary in complexity from black-box estimates of maximum flood peaks to detailed reconstruction of hydrographs. The methods may be divided into three categories:

- direct flow data;
- run-off modelling;
- regionalized flood formulae.

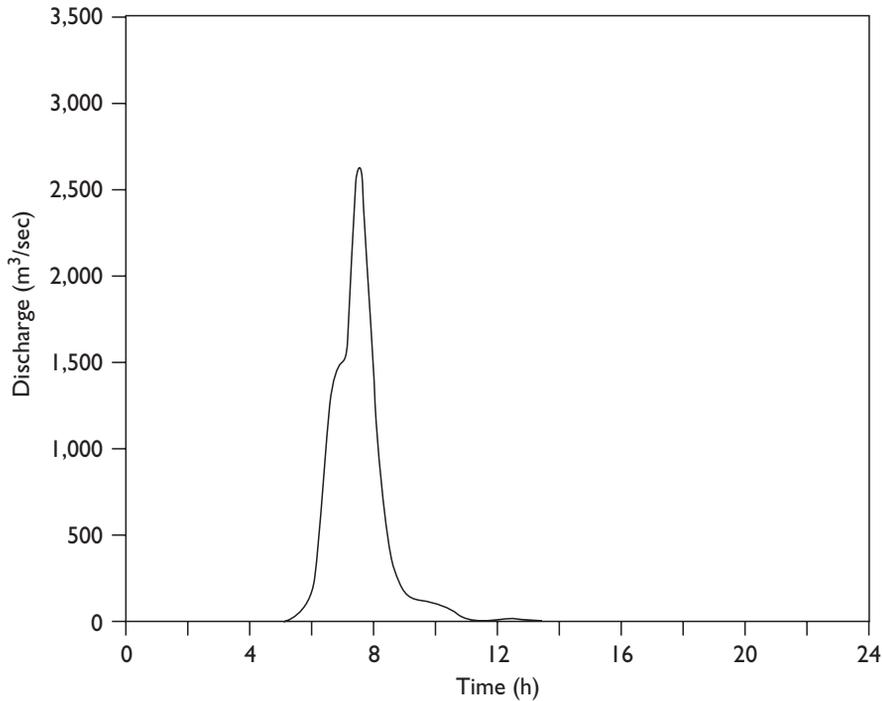


Figure 10.5 A flood hydrograph combining flows from several mountain catchments.

The three methods are listed in order of decreasing reliability and also decreasing data requirements. In all cases of flood assessment, estimates must be checked against historic evidence, local experience and practice, research and earlier studies. In spite of considerable efforts, flood estimation is still an uncertain science, and allowances must be made accordingly when engineering works are being designed.

10.3.2 Information surveys

To estimate rainfall run-off and select the design discharge (design flood), data on rainfall run-off, catchment and river characteristics need to be collected to an extent compatible with the project scope. During the data collection, it will become evident which flood estimation method is feasible to apply.

The information and hydrological data that ideally have to be collected are as follows:

- Topographic maps, with scale depending on project scope, but 1:25,000 or 1:50,000 will usually be adequate for identification of the geometric parameters of the catchments.
- Air photos or photo mosaics from which land-use can be studied when land-use maps are not available.
- Soil and vegetation maps or general descriptions.
- Water use in the project area, dams, reservoirs, abstractions for irrigation and other factors that may affect the run-off pattern.

Data
collection

Data needs

- Rainfall data, in terms of maximum intensities, as well as general climatologic information.
- Discharge gauging station data, in terms of annual flood discharges, and relation to stage discharge for the purpose of evaluating the accuracy of the flood flow data.

Field information Further data collection will have to take place in the field. At stream sites where bridges or major culverts are to be constructed, data on the following should be collected:

- cross-section area;
- bed material (in order to estimate the ‘manning roughness’ – see Section 10.3.3);
- longitudinal slope of stream bed;
- details of historic floods obtained from local residents (to estimate peak flood discharges).

River dynamics Possible future changes in the run-off regime may need to be evaluated, so evidence should be noted of morphological changes, building up or degradation of river bed levels, erosion patterns, debris size and controls, hydraulic controls, existing drainage structures, evidence of scour, stability of river banks, and land-use changes.

10.3.3 Direct flow data

Stream flow measurements The most common form of stream flow measurement comprises monitoring of the water levels at a suitable cross-section. The relationship between water level and discharge at the location is established through a series of discharge measurements (stage–discharge curve). Extrapolations of this curve will give estimates of the peak flood discharge for different flood water levels. An extreme value analysis of the annual maximum series will yield estimates of floods for different recurrence intervals. It is important to distinguish between the momentary peak and the maximum discharge averaged over a day, which is the discharge often quoted in hydrological yearbooks. For smaller catchments, there would be a significant difference between these values.

Slope-area method Discharge gauging stations are seldom situated close to the location of interest. Flood estimates therefore have to be made from levelling of marks left on vegetation or on the banks, calculation of the cross-sectional area, measurement of the slope of the stream bed and an estimate of the ‘manning roughness’. These observations and estimates combine to yield the flood discharge using the ‘manning equation’:

$$Q = A \times \frac{1}{n} \times R^{2/3} \times S^{1/2}$$

where Q is the discharge in m^3/sec ; A the cross-section area in m^2 ; n the manning roughness coefficient; and S the longitudinal slope of the stream bed. This method for estimating a peak discharge is termed the ‘slope-area method’.

Manning roughness coefficient By using the manning equation, direct estimates of the manning roughness coefficient over a fixed longitudinal section can be obtained by measuring discharge and area in two cross-sections and the longitudinal slope between them. This method provides a good estimate of the manning roughness coefficient. However, the

roughness changes with variation in factors such as vegetation and water level. Thus, large seasonal changes in the roughness must be expected, and the choice of manning roughness coefficient should reflect the flooding situation.

Another way of assessing flood estimates, when no gauging station is available at the location of interest, is by carrying out a series of discharge measurements to obtain the short-term statistical distribution. Simultaneous discharge measurements at an old discharge gauging station will provide short-term statistics for the same period, and the correlation between the two locations can be determined. Long-term data will be available from the gauging station, and it is possible to construct a long-term statistical distribution, which can then be transformed to the location of interest provided that a correlation exists.

Correlating
short- and
long-term data

10.3.4 Run-off modelling

The most widely used rainfall run-off relationship for ungauged areas is the 'rational method'. The method applies constant rainfall over the entire catchment and is thus most suitable for small catchments of sizes up to, say, a few square kilometres. The consequence of applying the rational method to larger catchments is an over-estimate of the discharges and, thus, a conservative design. The areal reduction factor (ARF) can be included to account for spatial variability over the catchment to compensate for the over-estimates. The basic form of the equation is

Rational
method

$$Q = \frac{C \times I \times A}{3.6}$$

where Q is the flood peak discharge at catchment exit (m^3/sec); C the rational run-off coefficient; I the average rainfall intensity over the whole catchment (mm/h) for a duration corresponding to the time of concentration; and A the catchment area (km^2).

The time of concentration is defined as the time required for the surface run-off from the remotest part of the drainage catchment to reach the location being considered. The time of concentration can be calculated by the 'kirpich formula':

Time of
concentration

$$T_c = \left(\frac{0.87 \times L^2}{1000 \times S} \right)^{0.385}$$

where T_c is the time of concentration in hours; L the length of main stream (km); and S the average slope of main stream (m/m).

Having determined the time of concentration, the corresponding rainfall intensity can then be obtained from the intensity-duration curve for the selected recurrence interval or return period.

The run-off coefficient, C , combines many factors influencing the rainfall run-off relationship, that is, topography, soil permeability, vegetation cover and land use. The run-off coefficient can be estimated using Table 10.1. The selection of the correct value for the run-off coefficient presents some difficulty because of the wide range of parameters reflected. Therefore, the value can vary from one time to another depending on changes especially in soil moisture conditions. Thus, the choice of run-off coefficient should be accompanied by field observations and by an evaluation of sensitivities of discharge estimates.

Run-off
coefficient

Table 10.1 Run-off coefficient for the rational method

C_T (topography)		C_S (soils)		C_V (vegetation)	
Very flat (<1%)	0.03	Sand and gravel	0.04	Forest	0.04
Undulating (1–10%)	0.08	Sandy clays	0.08	Farmland	0.11
Hilly (10–20%)	0.16	Clay and loam	0.16	Grassland	0.21
Mountainous (>20%)	0.26	Sheet rock	0.26	No vegetation	0.28
Run-off coefficient $C = C_T + C_S + C_V$					

Catchment area Catchment areas are determined from regional maps or aerial photographs. Varying topography and widely spaced contours on large-scale mapping sometimes makes it difficult to fix catchment boundaries.

Assumptions The rational method assumes that

- the design storm produces a uniform rainfall intensity over the entire catchment;
- the relationship between rainfall intensity and rate of run-off is constant for a particular catchment;
- the flood peak at the catchment exit occurs at the time when the whole catchment contributes to the discharge;
- the coefficient, C , is constant and independent of rainfall intensity.

Flood models More detailed run-off modelling can be carried out by taking into account all of the most significant factors, such as area, topography, land-use, vegetation, soil type, rainfall intensity, and aridity. One such model is termed the ‘generalized tropical flood model’ (Watkins and Fiddles 1984). In some cases, regionalized flood formulae have been established using regression analyses of flood and other catchment parameters. Although these may be useful at specific locations, they are seldom valid in other environments.

10.4 Hydraulic design

10.4.1 Design standards

Selecting recurrence interval The design discharge for a drainage facility depends on the selected flood frequency or recurrence interval. By selecting a large recurrence interval and a correspondingly large design flood, the probability of having such a flood occurring and the risk of damage is reduced, but costs of structures to accommodate such large floods are increased. Conversely, the selection of a small design flood reduces the initial cost of structures and increases the risk of damage from larger floods. When a recurrence interval is selected for a particular location, the designer is implying that the estimated effects of a larger flood on life, property, traffic and the environment do not justify constructing a larger structure at the time. Damage from lesser or more frequent floods should be minimal and acceptable.

Recommendations Different design standards will normally apply to different classes of road and type of drainage component being considered. Design manuals will often contain

Table 10.2 Recommended design standards for road classes and drainage components

	<i>Recurrence interval (years)</i>
<i>Road class</i>	
Expressways	100
Arterial roads	50
Collector roads	50
Access roads	25
<i>Drainage structure type (independent of road class)</i>	
Bridge	50
Large culverts (>1.2m diameter) and multiple culverts (>0.9 m diameter)	25
Small culverts (i.e. all other)	10
Embankments	10
Ditches and road surfaces	2

recommended design flood recurrence intervals. Table 10.2 recommends minimum design standards.

It is sound practice to supplement design calculations with sensitivity analyses. These can show consequences, such as damage for different recurrence intervals. Sensitivity analysis can be used to investigate adopting higher standards on the entire route, or only on those parts where there would be severe consequences. The analysis can also indicate where it might be reasonable to accept slightly lower standards, if the cost of achieving the preferred standards is too great.

Sensitivity
analysis

10.4.2 Risk

Risk can be defined in terms of the recurrence interval, or the probability of a flood of stated magnitude, being exceeded in any one year. Drainage structures are designed to accommodate a flow of a given return period or probability of occurrence. For example, for rural roads in the tropics, it is usual to design small culverts for a 1 in 10 years (or 10 per cent probability of occurrence event), but bridges for 1 in 50 or 100 years (or 2 per cent, or 1 per cent, probability of occurrence respectively). The lower standards are adopted for culverts because these are smaller structures and cheaper to repair if damage occurs from an over-design event. To build culverts to accommodate longer return period events, would increase capital costs beyond economic levels. The selection of the design flood recurrence interval involves an evaluation of the risk of disruption or damage to the road; that is possible loss of life, property damage, the interruption of traffic, and the economic consequences. This should be balanced with the capital and operational cost of providing and maintaining drainage. However, a risk-based approach to design can still present problems. For example, the 50-year flood can be exceeded at one location with minimal damage, but occurrence of the same frequency flood elsewhere might be disastrous.

Risk-based
design

Table 10.3 Probability of design standards being exceeded during design life

Design life 'L' years	Recurrence interval 'T' years and probability of design flow being exceeded (%)			
	1 in 10 years (10% probability)	1 in 20 years (5% probability)	1 in 50 years (2% probability)	1 in 100 years (1% probability)
20	88	64	33	18
30	96	78	45	26
50	99	92	64	39

Occurrence of over-design events

Over any period of time, there is a given probability that the design flow will be equalled or exceeded during the lifetime of the structure. On the basis that the original drainage design was to the standards quoted here the probability that drainage capacity has been exceeded can be calculated as

$$p = 1 - \left(1 - \frac{1}{T}\right)^L$$

where p is the probability; T the recurrence interval or return period; and L the lifetime.

Table 10.3 presents the probabilities of exceeding design standards for different design lives and recurrence intervals. It can be seen that, for example, there is a 96 per cent probability that a 30-year-old small culvert will have experienced a flow equal to or exceeding a designed capacity of 1 in 10 years since the road was built. For bridges, the corresponding probability is 45 per cent. The design life commonly used for a road is 20 years, and for a bridge is 50 years. Thus, there is a high probability that the design event will occur during the lifetime. Incidences of flooding for more frequent events will not happen simultaneously at every location on the road, but will occur randomly both temporally and spatially, dependent on the distribution of flood generating rainfall.

Climate change

The standard approach to drainage design using historical records does not take into account possible effects of climate changes. The statistical population may be changing and it is, therefore, unreliable to predict future probability of storm events based on historical records. Thus, drainage provision that was satisfactory in the past may not be adequate in the future. Unfortunately, it is not possible to predict the local effects of climate change, as the variation is generally not statistically significant. However, there is strong evidence that climate change is taking place, and there is perceived to be an increased incidence of flooding worldwide.

10.5 Longitudinal drainage components

Purpose

Longitudinal drainage conveys storm water from the surface to control the level of the water table in the subgrade, and to intercept groundwater and surface water flowing towards the road.

Crossfall

The road surface must be constructed with a sufficient camber or crossfall to shed rainwater quickly, and the formation of the road must be raised above the level of the local groundwater table. Wider pavements increase the catchment area, thus

increasing the quantity of stormwater that has to be removed. Flatter gradients, both transverse and longitudinal, increase water depth on the surface.

It is more important to maintain a minimum longitudinal gradient on kerbed than on unkerbed pavements to avoid stormwater spreading over the pavement. However, vegetation along the pavement edge may impede the run-off of water if the gradient is flat. Where the longitudinal gradient of the road is close to zero, the depth of side drains may have to be varied to obtain sufficient gradient of the ditch. Longitudinal gradients should preferably be not be less than 0.3 per cent for kerbed pavements, and not less than 0.2 per cent in very flat terrain. A minimum longitudinal gradient of 0.3 per cent should be maintained within 15 m of the bottom of sag curves on kerbed pavement. Zero gradients and sag vertical curves should be avoided on bridges.

Longitudinal
gradients

Ideally, the base and sub-base should extend below the shoulder to the side ditches. When ditches are lined with concrete or masonry, drainage outlet pipes or weep holes must be provided through the lining.

Sub-surface
drainage

If it is too costly to extend the base and sub-base material below the shoulder, drains (grips) at 3–5 m intervals should be cut through the shoulder to a depth of 50 mm below sub-base level. Grips should be backfilled with base material, or more permeable material, and should be given a fall of 1 in 10 to the side ditch. Alternatively, a continuous drainage layer 75–100 mm thickness of pervious material can be laid under the shoulder at the level of the underside of the sub-base. Perforated drainpipes can also be used to drain the road pavement.

Grips

Ditches, gutters, turnouts, chutes, and intercepting ditches may be used to provide open road-side drainage.

Road-side
drainage
Ditches

Ditches are channels provided to remove the run-off from the road pavement, shoulders, and cut and fill slopes. The depth of the ditch should be sufficient to remove the water without risk of saturating the subgrade. Ditches may be lined to control erosion. Unlined ditches should preferably have side slopes not steeper than 4 to 1 horizontal to vertical.

Gutters are channels at the edges of the pavement or the shoulder formed by a kerb or shallow depression. Gutters are paved with concrete, brick, stone blocks, or some other structural material. Spacing between outlets on kerbed road sections depends on run-off, longitudinal gradient and permissible water depth along the kerb.

Gutters

Turnouts or mitre drains are short, open, skew ditches used to remove water from the road-side ditches or gutters. Use of turnouts reduces the necessary size of the side ditches, minimizes the velocity of water and thereby the risk of erosion. The interval between turnouts depends on run-off, permissible velocity of water and slope of the terrain. To prevent the flow through turnouts from generating soil erosion at the outlet, the discharge end of the turnout should be fanned out.

Turnouts

Chutes are open, lined channels or closed pipes, used to convey water from gutters and side ditches down fill slopes, and from intercepting ditches down cut slopes. On long slopes, closed (pipe) chutes are generally preferable to open chutes. The inlet of chutes must be designed to prevent water bypassing the chute and eroding the slope. The outlet must also be designed to prevent erosion at the outlet (see Figure 10.6). The distance between chutes will depend on the capacity of gutters or ditches.

Chutes

Intercepting (or cut-off) ditches are located on the natural ground near the top edge of a cut slope to intercept the run-off from a hillside before it reaches the road.

Intercepting
ditches

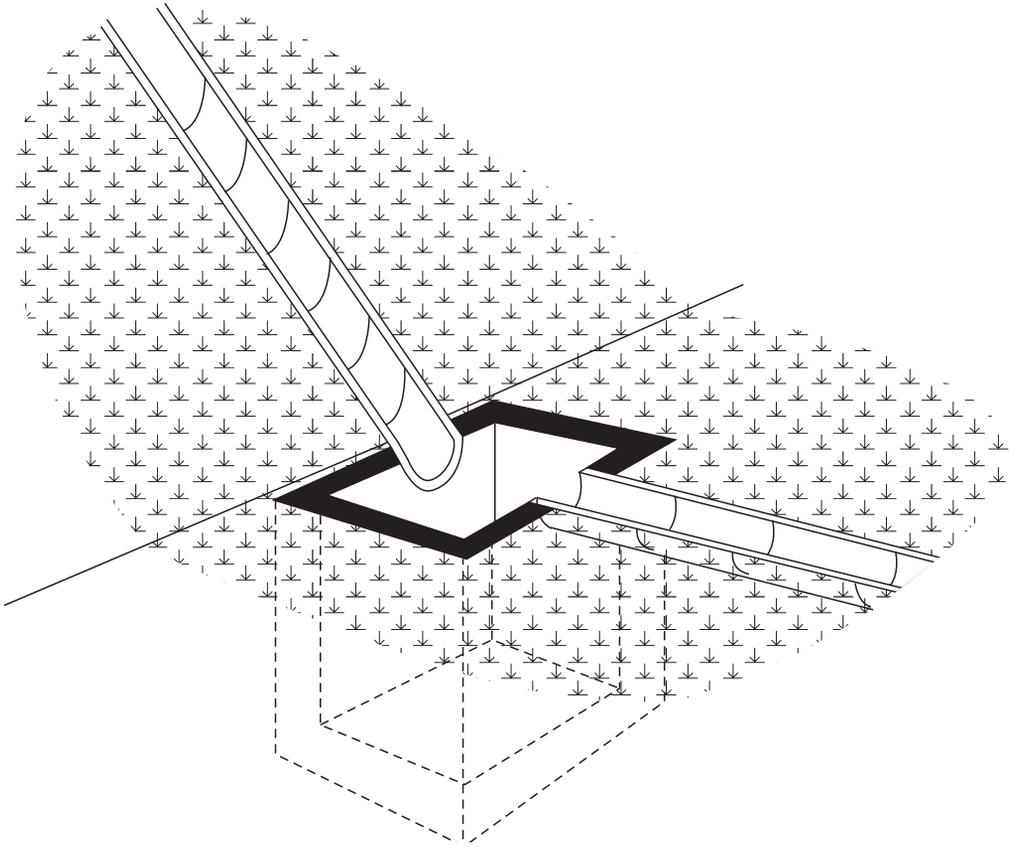


Figure 10.6 Chute outlet.

Intercepting the surface flow reduces erosion of cut slopes and roadside ditches, lessens silt deposition and infiltration in the area of the road, and decreases the likelihood of flooding the road in severe storms. Intercepting surface water is particularly important in arid and semi-arid regions because of generally low water infiltration capacity and high tendency to erosion of arid soils. Intercepting ditches may be built at least 3 m from the top of the cut slope and, generally, should have a flat grade until the water can be spread or emptied into a natural watercourse. However, because they are distant from the road, the maintenance of intercepting ditches tends to be neglected. Poorly maintained interceptors can result in slip failures in cutting slopes.

Hydraulic calculation

Most longitudinal drains are open channels. Their hydraulic capacity is often designed to contain a 5 or 10-year frequency storm run-off. The estimated run-off for the 2-year frequency storm can be used for determining the needs, type, and dimensions of special channel linings for erosion control. The design discharge is calculated according to the rational formula, and the capacity of an open channel is calculated according to the manning equation. This gives a reliable estimate of uniform flow conditions:

$$Q = A \times v = A \times \frac{1}{n} \times R^{2/3} \times I^{1/2}$$

Table 10.4 Manning roughness coefficient

Type of lining	Water depth		
	0–150 mm	150–600 mm	>600 mm
Concrete	0.015	0.013	0.013
Grouted rip-rap	0.040	0.030	0.028
Stone masonry	0.042	0.032	0.020
Soil cement	0.025	0.022	0.020
Asphalt	0.018	0.016	0.016
Bare soil	0.023	0.020	0.020
Rock cut	0.045	0.035	0.025
Rip-rap (25 mm)	0.044	0.033	0.030
Rip-rap (50 mm)	0.066	0.041	0.034
Rip-rap (150 mm)	0.104	0.069	0.035
Rip-rap (300 mm)	—	0.078	0.040

Source: FWHA (1988).

where Q is the capacity in m^3/sec ; A the channel cross-section area in m^2 ; v the mean velocity in m/sec ; n the Manning roughness coefficient; R the hydraulic radius A/P in m ; I the slope in m/m ; and P the wetted perimeter in m ; Table 10.4 indicates some Manning roughness coefficients at different water-depths.

10.6 Cross drainage components

Cross drainage serves to convey water across the alignment of the road. Cross drainage structures can be very costly and it is, therefore, important to analyse all major cross drainage along the road before final selection of a new road alignment. Where there is a choice in the selection of the position of a stream crossing, it is desirable that it should be located:

Road alignment and drainage

- on a straight reach of the stream, away from bends;
- as far as possible from the influence of large tributaries;
- on a reach with well-defined banks;
- at a site which makes straight approach roads feasible;
- at a site which makes a right-angle crossing possible.

To determine the type of cross drainage, relevant information on hydrology must be collected, and predictions are needed about the level of traffic.

The following types of structure may be considered:

Type of structure

- fords
- drifts
- culverts
- bridges.

The simplest river crossing is a ford. This utilizes a suitable existing river-bed and is appropriate for shallow, slow-moving water courses. They are appropriate for traffic

Fords

volumes up to about 100 vehicles per day. Gravel or stones can be used to line the bottom of the ford to provide a firm footing for vehicles. Fords should normally only be used for rivers that are not subject to flash floods, as they may cause the ford to be washed away. However, repair or replacement is cheap.

Drifts

If the river-bed is not able to carry vehicles, a concrete slab may be constructed in the bed. This type of crossing is known as a 'drift' (see Figure 10.7). A drift is suitable as a crossing for rivers that are normally fordable, but prone to floods. Drifts are appropriate for similar traffic volumes to fords.

Vented drift

Where the river is running most of the year, the drift can be provided with openings to permit water to pass below road level, and reduce the frequency and

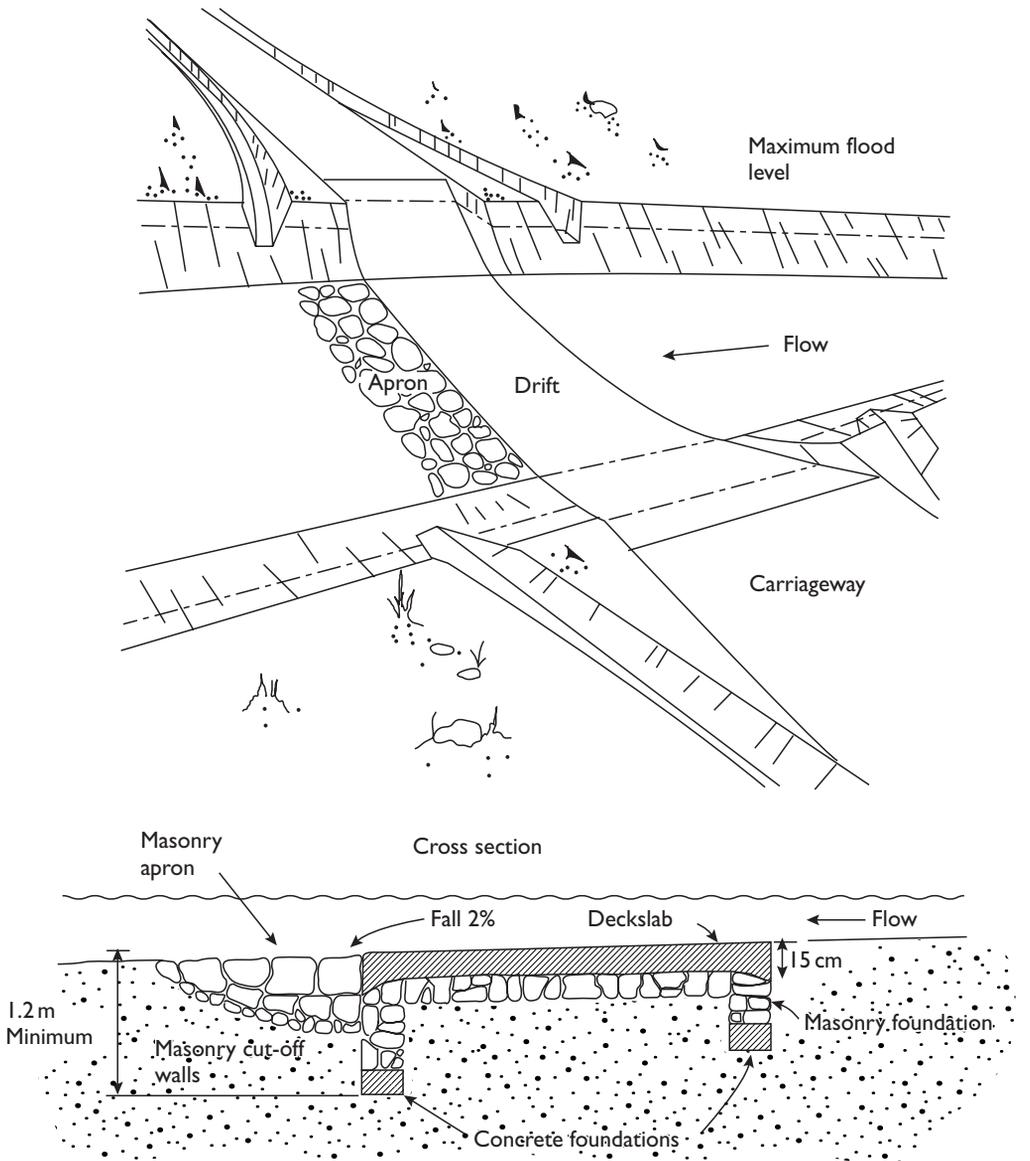


Figure 10.7 Drift.

Source: Antoniou et al. (1990).

depth of over-topping during flash floods. This type of crossing is called a ‘vented drift’ or a ‘causeway’.

Culverts are used to convey water from streams crossing the road and to carry water from one side-ditch to the other. The difference between culverts and bridges is, that culverts are placed in the embankment below the road pavement, whereas bridge decks form part of the pavement. Usually bridges have a larger span than culverts. Bridges may also be constructed to allow for the passage of boats, whereas culverts are frequently designed for full-flow through the cross-section. Culverts usually consist of a concrete or steel pipe (see Figure 10.8), or a reinforced concrete box. Most countries make precast concrete pipes of up to 1-m diameter and these may be cost-effective, provided that they can be transported and handled. Corrugated galvanized steel pipes, often known by the trade name ‘Armco’, are available in large diameters, and are lighter and easier to handle. There should be little maintenance required for either material, other than an annual inspection and clearing of accumulated silt or debris, although corrosion may occur to metal pipes in some circumstances.

Most culverts have an upstream headwall and terminate downstream with an endwall. Headwalls direct the flow into the culvert, while endwalls provide a transition from the culvert to the outlet channel. Both protect the embankment from erosion by flood water. Headwalls and endwalls are sometimes both called ‘headwalls’. Straight headwalls placed parallel to the roadway are used mainly with smaller pipe culverts. For larger culverts, headwalls are normally supplemented with wingwalls at an angle to the embankment (see Figure 10.9). Most headwalls and wingwalls are concrete, although masonry is also used. In all cases, a cut-off wall (toe-wall) extending below the level of expected scour should be incorporated in the design of the outlet. Often, a paved apron extending beyond the cut-off wall, is a wise addition. The culvert inlet may also be provided with an apron (see Figure 10.8).

For larger volumes of water, several pipes can be used in parallel under the road. Multiple pipes can also be used where the planned embankment height is insufficient to cover a single pipe of sufficient diameter. However, pipes of less than 1-m diameter are not recommended since they are difficult to maintain.

Culverts

Inlets and outlets

Multiple culverts

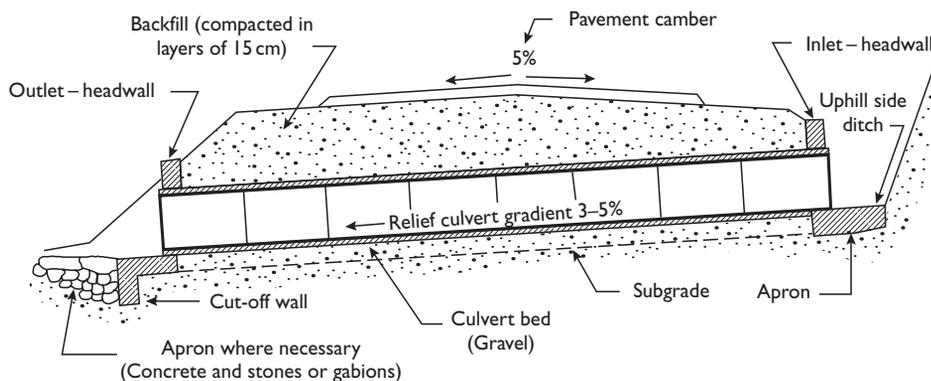


Figure 10.8 Pipe culvert.

Source: Beenhakker et al. (1987).

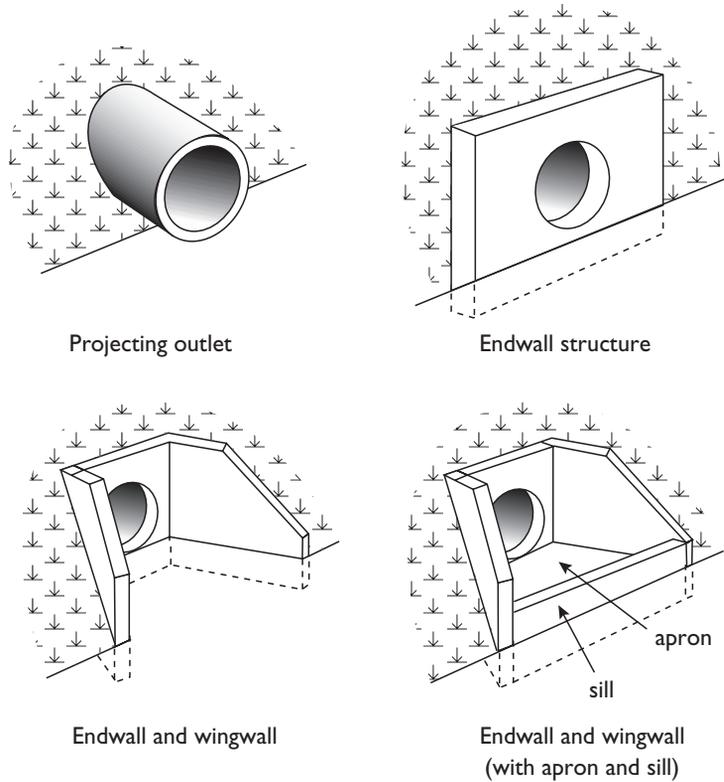


Figure 10.9 Types of culvert outlets.

Box culverts

Reinforced-concrete box culverts may also be used either singly or in parallel where relatively large volumes of water are expected. These culverts are normally cast in place, although smaller sizes may be precast.

Bridges

Bridges are required for crossing streams and rivers where culverts would provide insufficient capacity, or where the road crosses an obstruction such as a railway or canal. There is no restriction to traffic unless the width of the structure is less than the road width.

10.7 Culvert design

Design discharge

The design discharge used for culverts is usually estimated on the basis of a preselected recurrence interval. The culvert is designed to operate in a manner that is within acceptable limits of risk at that flow rate.

Location

Location of culverts should be selected carefully. The alignment of a culvert should generally conform to the alignment of the natural stream. The culvert should, if possible, cross the road at a right-angle to minimize cost. However, skew culverts located at an angle to the centre-line of the road are needed in many instances. The slope of the culvert should generally conform to the existing slope of the stream. To avoid silting, the slope of the culvert should not be less than 1 per cent.

Design factors

The following concepts influence the determination of culvert size:

- headwater depth;
- tailwater depth;

- outlet velocity;
- culvert flow with ‘inlet control’;
- culvert flow with ‘outlet control’.

Culverts generally constrict the natural flow of a stream and cause a rise in the upstream water level. The height of the water at the culvert entrance is termed headwater elevation, and the total flow depth in the stream measured from the culvert inlet invert is termed headwater depth. The headwater elevation should not be too high. High headwater may damage the culvert and the road, and interrupt the traffic. High headwater may also damage upstream property and cause hazards to human life. Headwater depth is a function of the discharge, the culvert size and the inlet configuration. The headwater elevation for the design discharge should be at least 500 mm below the edge of the shoulder elevation. A ratio between headwater depth and height of culvert opening equal to 1.2 is recommended in cases where insufficient data are available to predict the flooding effect from high headwater.

Headwater

Tailwater depth is the depth of flow in the downstream channel measured from the invert at the culvert outlet. Tailwater depth can be an important factor in culvert hydraulic design, because a submerged outlet may cause the culvert to flow full rather than partially full. An approximation of the tailwater depth can be made using the Manning equation, if the outlet channel is reasonably uniform in cross-section, slope and roughness. However, tailwater conditions during floods are sometimes controlled by downstream obstructions or by water conditions. A field inspection should always be made to check on features that may influence tailwater conditions, but considerable experience is then needed to predict tailwater depth.

Tailwater

The outlet velocity, measured at the downstream end of the culvert, is usually higher than the maximum natural stream velocity. This higher velocity can cause stream bed scour and bank erosion for a limited distance downstream.

Outlet
velocity

There are two major types of culvert flow:

Culvert flow

- flow with inlet control;
- flow with outlet control.

For each type of control, a different combination of factors control the hydraulic capacity of a culvert. The determination of actual flow conditions can be difficult. Therefore, the designer should check for both types of flow and design for the most adverse condition.

A culvert operates with inlet control when the flow capacity is controlled at the entrance by the following factors:

Inlet control

- culvert type (shape of barrel);
- type of culvert inlet;
- culvert cross-sectional area;
- headwater depth.

When a given culvert operates under inlet control, the headwater depth determines the culvert capacity, with the barrel usually flowing only partially full. Figure 10.10 illustrates the inlet control flow for unsubmerged and submerged entrances. For culverts flowing under inlet control, wingwalls improve the hydraulic characteristics. By

rounding the entrances of circular pipes, or tapering the inlets of rectangular or square cross-sections, the flow accommodated at a given head can be substantially increased.

Outlet control When a culvert operates under outlet control, the flow capacity is determined by the same factors as under inlet control but, in addition, the performance depends on

- roughness of the inner surface of the culvert (barrel roughness);
- longitudinal slope of the culvert (barrel slope);
- tailwater depth or critical depth.

Culverts operating under outlet control may flow full or partly full, depending on various combinations of the determining factors. Typical outlet control flow situations are shown in Figure 10.11.

Energy conservation The design procedure is based on a computation using the principle of ‘conservation of energy’:

$$\left(z + \frac{p}{\gamma} + \frac{\alpha \cdot v^2}{2g} \right)_A = \left(z + \frac{p}{\gamma} + \frac{\alpha \cdot v^2}{2g} \right)_B + \Delta H_{AB}$$

where z is the elevation in m; p the pressure; v the velocity (m/sec); γ the specific gravity in kN/m^3 ($\gamma = \rho g$); α the velocity distribution coefficient, ($\alpha = 1.1$ in turbulent flow); and ΔH_{AB} the energy loss between section A and B.

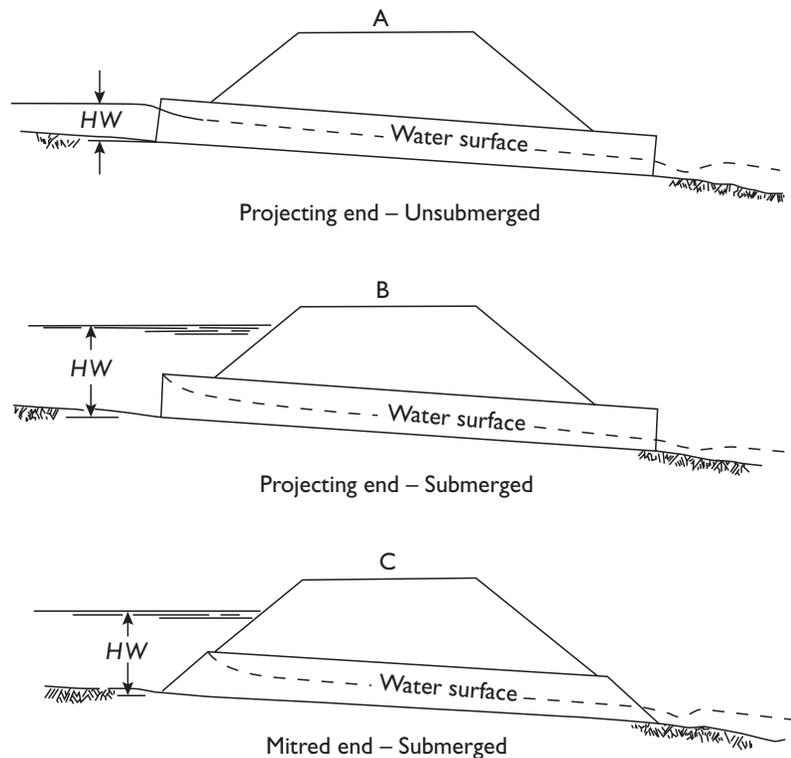


Figure 10.10 Inlet control.

Source: Transportation Research Board (1978).

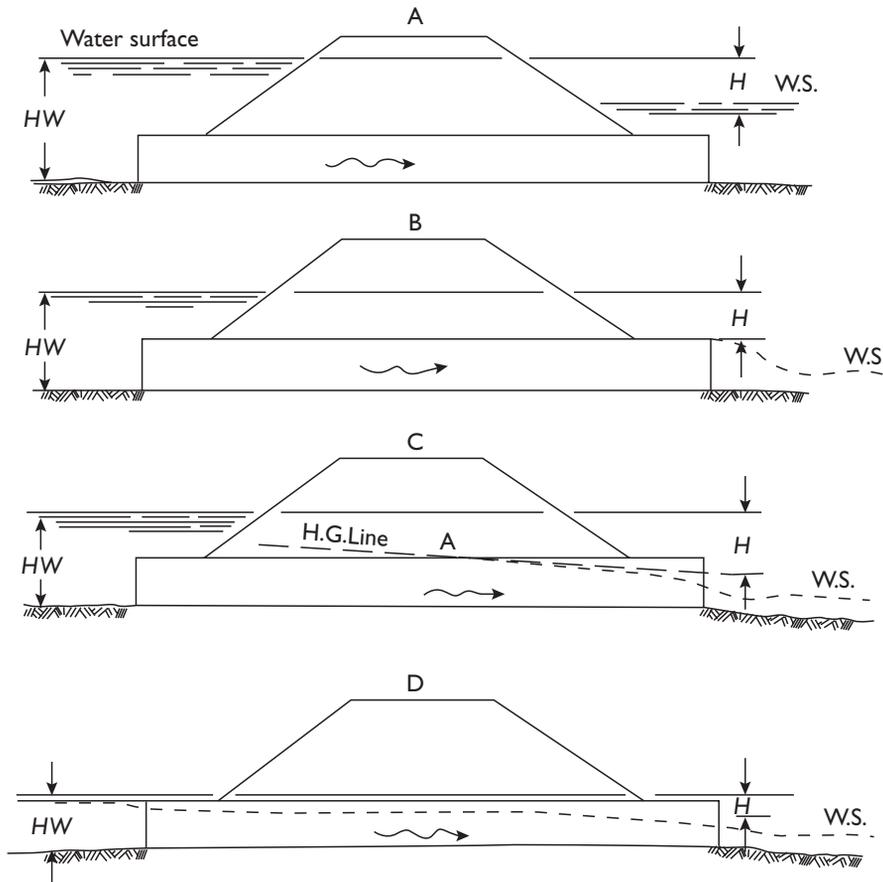


Figure 10.11 Outlet control.

Source: Transportation Research Board (1978).

The energy equation takes into account energy losses in the inlet and outlet, as well as energy loss due to friction in the culverts. Discharge capacities for different culvert designs and associated water levels can be calculated directly, corresponding to the actual flow conditions.

10.8 Erosion and scour protection

10.8.1 Ditches

Usually ditches are unlined, and often suffer from scour. Tables 10.5 and 10.6 indicate maximum permissible velocities (V_p) that could be used without scour for maximum water depths of 1 m.

The amount of erosion control and maintenance can be minimized to a great extent by using the following:

- flat side slopes;
- turnouts and intercepting ditches;
- ditch checks or drop structures (see Figure 10.12);

Permissible
velocity

Erosion
control

Table 10.5 Permissible velocities for ditches lined with various grass covers

Type of lining	Permissible velocity (m/sec)
Well-established grass on any good soil	1.8
Meadow type grass, for example, blue grass	1.5
Bunched grasses, with exposed soil between plants	1.1
Grains and stiff stemmed grasses that do not bend over under shallow flow	0.9

Source: MoWC (1982).

Table 10.6 Permissible velocities in excavated ditches

Soil type or lining (no vegetation)	Permissible velocity (m/sec)		
	Clear water	Water-carrying fine silts	Water-carrying sand and gravel
Fine sand (non-colloidal)	0.45	0.75	0.45
Sandy loam (non-colloidal)	0.55	0.75	0.60
Silt loam (non-colloidal)	0.60	0.90	0.60
Ordinary firm loam	0.75	1.05	0.70
Volcanic ash	0.75	1.05	0.60
Fine gravel	0.75	1.50	1.15
Stiff clay (very colloidal)	1.15	1.50	0.90
Graded, loam to cobbles (non-colloidal)	1.15	1.50	1.50
Graded, silt to cobbles (colloidal)	1.20	1.70	1.50
Alluvial silts (non-colloidal)	0.60	1.05	0.60
Alluvial silts (colloidal)	1.15	1.50	0.90
Coarse gravel (non-colloidal)	1.20	1.85	2.00
Cobbles and shingles	1.50	1.70	2.00
Shales and hard pans	1.85	1.85	1.50
Rock	Negligible scour at all velocities		

Source: MoWC (1982).

- asphalt or concrete kerbs, with chutes on high embankments;
- protective lining of ditches by rigid linings, such as cast-in-place concrete, stone masonry or grouted rip-rap;
- protective covering of ditches by flexible linings, such as rip-rap, wire-enclosed rip-rap, vegetation or synthetic material.

Rigid linings

Rigid linings are useful in flow zones where shear stress is high or where non-uniform flow conditions exist, such as at transitions in ditch shape or at an energy dissipation structure. In areas where loss of water or seepage from the ditch is undesirable, they provide an impermeable lining. Cast-in-place concrete or masonry linings often break up and deteriorate if foundation conditions are poor. Once a rigid lining deteriorates, it is very susceptible to erosion.

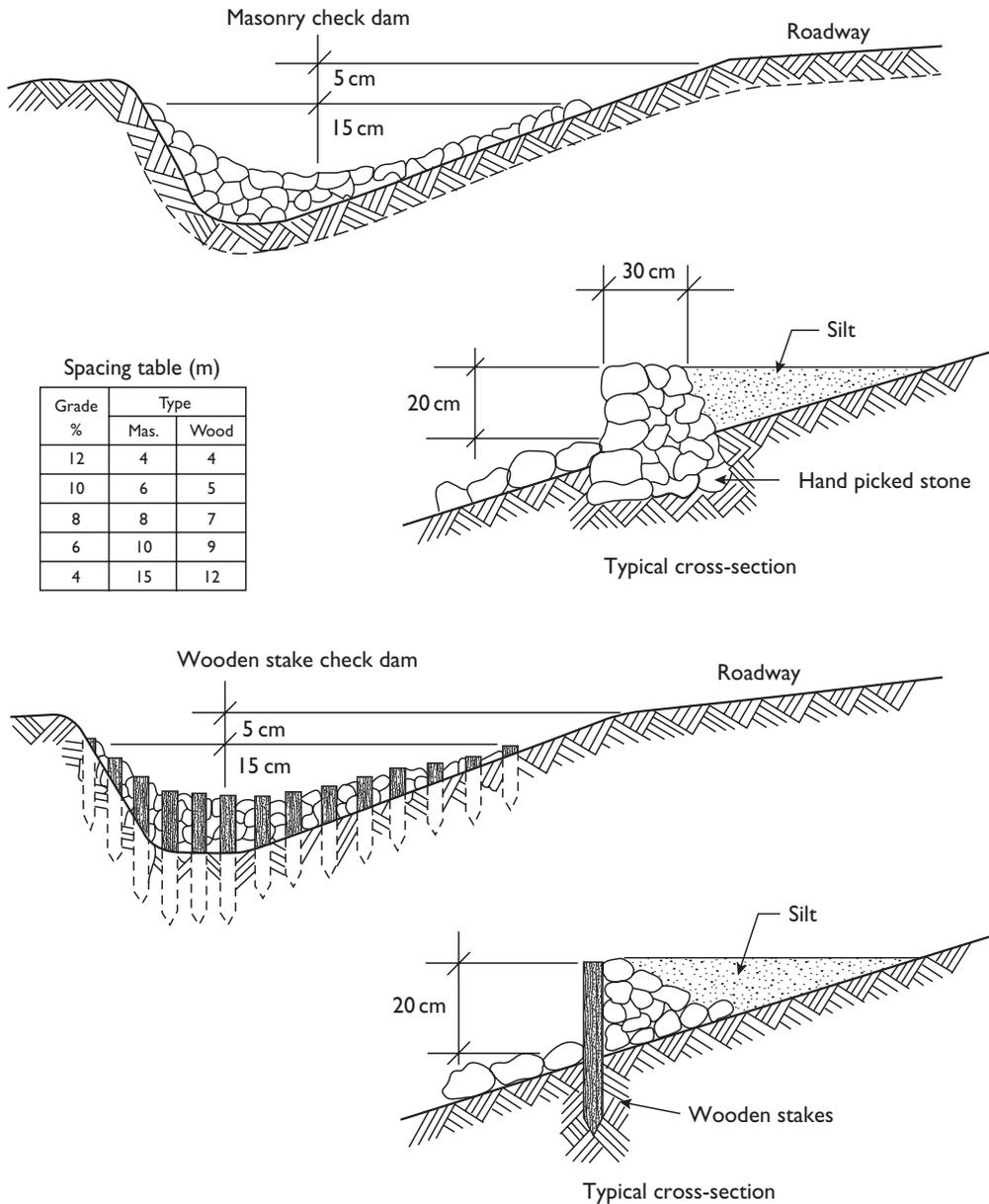


Figure 10.12 Ditch checks.

Source: Beenhakker et al. (1987).

Flexible linings are suitable for hydraulic conditions similar to those requiring rigid linings. However, because flexible linings are permeable, protection of the underlying soil may be required to prevent washouts. For example, filter cloth is often used with rip-rap to inhibit soil piping. Lining with vegetation is suited to hydraulic conditions where uniform flow exists and shear stresses are moderate. Flexible linings are not suited to continued flow conditions or long periods of being submerged.

Flexible linings

10.8.2 Culverts

Inlet	<p>Unchecked erosion is the prime cause of culvert failure. At culvert inlets, erosion from vortexes and flow over wingwalls is generally not a major problem. However, erosion may be caused by high inlet flow velocity. Velocity near the inlet may be estimated by dividing the flow rate by the area of the culvert opening. The risk of channel erosion should be judged on the basis of this approach velocity. The extent of the slope protection can be limited to cover only an area in the immediate vicinity of the culvert entrance. However, where the flow enters the transition between the embankment and the wingwalls, high local velocities and flow disturbances are expected, justifying the slope protection to be extended a certain distance beyond the wingwalls. Where water flows along the edge of an embankment before reaching the culvert entrance, velocities may be high. Embankments upstream of the culvert therefore need to be protected from erosion. Erosion protection should also be provided if flow velocity near the inlet indicates a possibility of scour threatening the stability of wingwall footings. A concrete apron with cut-off wall between wingwalls is the most satisfactory means for providing this.</p>
Outlet	<p>The greatest scour potential is at the culvert outlet, where high velocities may necessitate scour protection or energy dissipation. Thus, determination of the flow condition, scour potential and channel erodibility at the outlet should be standard procedure in the design of all road culverts. The only safe procedure is to design on the basis that erosion will occur at a culvert outlet and in the downstream channel. A reasonable procedure is to provide at least minimum protection, and then inspect the outlet channel after major rainfall to determine if the protection needs to be increased or extended. Two types of scour can occur in the vicinity of culvert outlets: local scour and general channel degradation.</p>
Local scour	<p>Local scour is the result of high-velocity flow at the culvert outlet, but its effect extends only a limited distance downstream. Natural channel velocities are almost always lower than culvert outlet velocities, because the channel cross-section and its flood plain are generally larger than the culvert flow area. Energy dissipaters, such as drop structures, rip-rap basins, stilling basins are the most appropriate structures to prevent scour at the outlet.</p>
Channel degradation	<p>Degradation may progress in a fairly uniform manner over a long length of the channel, or may consist of one or more abrupt drops (head-cutting), which progress upstream with every run-off event. The highest velocities will be produced by long, smooth-barrel culverts on steep slopes, and protection of the outlet channel at most sites is necessary. However, protection will also be required for culverts on mild slopes.</p>

10.8.3 Bridges

Types of scour	<p>Three principal forms of scour need to be considered when designing major drainage structures and bridges. Natural scour occurs in streams where there is migration of the bed, which shifts the direction of flow, and at bends and constrictions. General scour occurs due to the introduction of an obstruction in a stream channel, such as a bridge pier, since this increases velocities in the constricted section. General scour is avoided by designing bridges with a sufficiently large span. Local scour occurs in</p>
----------------	---

more concentrated locations, such as around bridge piers, because of high local velocities and flow disturbances, such as eddies and vortices. The effects of natural, general and local scour are generally additive where they occur at the same location. The total effect of scour is often estimated by calculating only the effect of local scour, combined with the use of a large safety factor.

Bridge piers built in alluvial material are often exposed to undercutting and may settle or collapse, unless set deeply, or provided with local protection. The depth of local scour that may be expected in any given location depends upon the duration and peak-flow of the flood, the susceptibility of the bed material to rapid erosion, the depth of flow, the mean flow velocity and the pier size. Even when field measurements taken during a flood are available, estimation of scour depth is difficult. A general and somewhat conservative rule-of-thumb is that maximum depth of scour is equal to twice the pile diameter.

Local scour

To evaluate whether sediment movements on the bed occur, it is necessary to determine whether the actual bed shear stresses, induced by the water, are larger than the resistance of the sediment. This requires determination of whether the mean velocity of the approach flow is larger than the critical velocity at which sediment starts to move. The critical velocity is assessed through the dimensionless critical bed shear stress, θ_c , which is obtained from a 'shields diagram', as in Figure 10.13. As input to the shields diagram, the boundary Reynolds number, describing the flow conditions, must be evaluated:

Shields parameter

$$R = \frac{\sqrt{g \cdot y \cdot I}}{\nu} \cdot D_{50}$$

where R is the boundary Reynold's number; g the acceleration of gravity (9.81 m/sec^2); I the hydraulic gradient or river slope in m/m ; ν the kinematic viscosity of water ($10^{-6} \text{ m}^2/\text{sec}$ at 20°C); and D_{50} the mean size of bed material (in m). Finally, the critical velocity is calculated from the following:

$$V_c^2 = \theta_c \times \left(6 + 2.5 \ln \frac{y}{2.5 D_{50}} \right)^2 \times \left(\frac{\gamma_s}{\gamma} - 1 \right) \times g \times D_{50}$$

where V_c is the critical velocity in m/sec ; θ_c the dimensionless critical bed shear stress; γ_s the specific weight of sediment grains in N/m^3 ; and γ the specific weight of water in N/m^3 .

Protection works, such as rock aprons, will limit the depth of scour. Rock aprons should be laid at least below the general scour level. River banks can be protected by rock aprons as well as by other revetments, such as gabions and precast concrete blocks. At bridge crossings, river training works can help control flow through the bridge openings and control erosion at river banks.

Protection

The best solution for minimizing scour and flood damage at new bridges is to

New bridges

- locate the bridge in a way that adverse flood flow patterns are avoided;
- streamline bridge elements to minimize obstructions to the flow;
- deepen the foundations to accommodate scour.

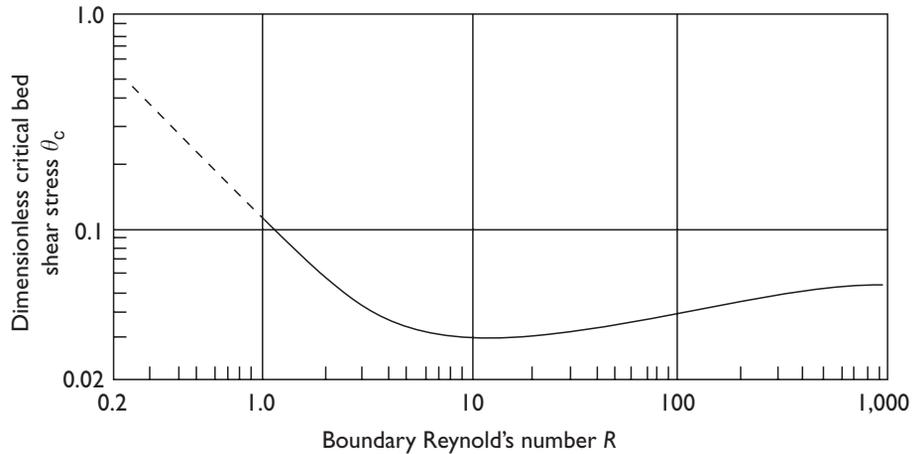


Figure 10.13 Shields diagram.

Source: ASCE (1975).

Existing bridges

For existing bridges, the alternatives available to protect the bridge from scour and flood damage are listed here in a rough order of cost:

- providing rip-rap at piers and abutments;
- constructing river training works;
- straightening out and widening the stream;
- strengthening the bridge foundations;
- constructing energy dissipation devices such as cills or drop structures;
- constructing relief bridges, or lengthening existing bridges;
- constructing a causeway (a road section where flood water can pass over the road) in the vicinity of the bridge by lowering the longitudinal profile of the road.

References

- Antoniou, J., Guthrie, P. and de Veen, J. (1990) *Building Roads by Hand: An Introduction to Labour-based Road Works*, Burnt Hill: Longman for the International Labour Office.
- ASCE (1975) 'Sedimentation engineering' *Manuals and Reports on Engineering Practice No. 54*, Washington, DC: American Society of Civil Engineers.
- Beenhakker, H.L. et al. (1987) *Rural Transport Services. A Guide to Their Planning and Implementation*, London: Intermediate Technology Publications.
- FWHA (1988) 'Design of roadside channels with flexible linings', *Federal Highway Administration Hydraulic Engineering Circular No. 15*, Washington, DC: US Department of Transportation.
- MoWC (1982) *Botswana Road Design Manual*, Gaborone: Ministry of Works and Communications.
- Transportation Research Board (1978) 'Compendium 3: Small drainage structures', *Transportation Technology Support for Developing Countries*, Washington, DC: National Research Council.
- Watkins, L.H. and Fiddes, D. (1984) *Road and Urban Hydrology in the Tropics*, London: Pentech Press.
- WMO (1970) *Guide to Hydrometeorological Practices*, Geneva: World Meteorological Organization.

Geometric design controls

Kent Falck-Jensen

11.1 Basic considerations

11.1.1 Current basis of standards

Historically, geometric design standards used in developing countries have been based on those in industrialized countries. Standards used by many developing and emerging countries are based on early American standards, with modifications to take into consideration the special conditions in the individual country. Many countries have developed road design manuals with guidelines on geometric design standards or regional road design standards like the *SATCC Recommendations on Road Design Standards* (SATCC Technical Unit 1995), which were prepared to harmonize standards and specifications for road design, construction and maintenance within the member states of the Southern Africa Transport and Communications Commission. These aim to provide a set of standards for the regional trunk road network that can be accepted by all member countries. The recommendations are a compromise between existing design standards in use within the region.

Basis of standards

A study carried out by the then Transport and Road Research Laboratory (Kosasih *et al.* 1987) considered the application of the American (AASHTO 1984), Australian (NAASRA 1980) and British (Department of Transport 1981) standards to developing countries. The study concluded that roads and driving conditions were sufficiently different in developing countries to question the application of industrialized standards in these situations. Further studies were then carried out (Boyce *et al.* 1988), and the resulting design guide was issued as *TRRL Overseas Road Note 6* (TRRL Overseas Unit 1988), which reflected differences in conditions in the following significant areas:

Application of standards

- Traffic mix
- Rate and nature of road accidents
- Level of economic development
- Maintenance capability.

11.1.2 Traffic mix

The needs of road users in developing and emerging countries are often very different from those in the industrialized countries. For example, pedestrians,

bicycles and animal drawn carts are often important components of the traffic mix, even on major roads. Trucks and buses, especially minibuses, often represent the largest proportion of the motorized traffic, while traffic composition in the industrialized countries is dominated by the passenger car. As a result, there may be less need for high-speed roads and it will often be more appropriate to provide wide and strong shoulders.

11.1.3 Rate and nature of road accidents

Accident rates Accident rates in developing and emerging countries are considerably higher than in industrialized countries for similar levels of vehicle flow and geometric design (see Chapter 4). Contributory factors are road user behaviour, vehicle condition and maintenance. Thus, from the point of view of safety, it appears that geometric standards used in the developed countries are not fully applicable to the developing world.

Relationship with road geometry The relationships between accident rates and road geometry suggest that the number of junctions per kilometre is the most significant factor, followed by horizontal and vertical curvature. High accident rates have been observed on gravel roads. Among the possible causes of this might be poor geometry, slipperiness of the surface in wet weather and poor visibility caused by dust and high vehicle speeds. Some studies have shown that accident rates decrease with reduced road roughness and it is likely that keeping gravel road surfaces well maintained or sealing surfaces could reduce accident rates.

Pedestrian accidents Pedestrian accident rates tend to be very high in developing and emerging countries. Often accidents are related to pedestrians choosing to cross at unsuitable locations or drivers ignoring speed restrictions. The design should therefore aim at reducing conflicts between the road users, or providing priorities in an obvious and positive way.

11.1.4 Economic development levels

Low traffic volumes As a result of the lower level of economic development, the traffic volumes on most rural roads in developing and emerging countries are relatively low. Thus, providing a road with high geometric standards will not normally be feasible, since transport cost savings may not offset construction costs. The provision of wide carriageways, flat gradients and full overtaking sight distance may therefore be inappropriate. Also, in countries with weak economies, design levels of comfort used in industrialized countries may well be an unaffordable luxury.

Stages of development When developing appropriate geometric design standards, the first step should normally be to identify the purpose of the road in question. It is convenient to define the purpose in terms of three distinct stages of development (McLean 1978), as follows.

Stage I: provision of access Initially, it is necessary to establish a road network to at least provide a basic means of access and communication between centres of population and between farms and markets. At this stage, little attention is paid to geometric standards as it is much more important to consider whether a road link exists at all or, if it does, whether it is 'passable' at all times. The choice of standards will be governed only

by issues such as traction requirements and turning circles. Absolute minimum standards can be used to provide an engineered road.

The next stage is to build capacity into the road network. Geometric standards probably have little to contribute to this except in the areas of road width and gradient. Some studies have suggested that, for relatively low traffic volumes, road width in excess of 5 m cannot be justified in terms of accident reductions or traffic operations (Boyce *et al.* 1988). However, a wider cross-section may be appropriate on sections where restrictions on sight distance apply and on sharp bends to enable the curves to be negotiated by the heaviest anticipated vehicle type. Much more important factors are whether or not a road is paved, or whether it has sufficient structural strength to carry the traffic wishing to use it.

This final stage is to consider operational efficiency of the traffic and it is at this stage that geometric standards become really important.

Developing and emerging countries, by their very nature, will not usually be at Stage 3 of this development sequence; indeed many will be at the first stage. However, design standards currently in use are generally developed for countries at Stage 3 and they have been developed for roads carrying relatively large volumes of traffic. For convenience, these same standards have traditionally been applied to low-volume roads, which has led to uneconomic and technically inappropriate designs.

11.1.5 Maintenance capabilities

The inability of many countries to finance adequate maintenance for their roads leads many road designers to adopt high standard and expensive design solutions assuming that, without appropriate maintenance, a high standard road will last longer than a poorer one. However, there is little evidence to suggest that a road designed to a high standard will suffer proportionally less than a road built to an appropriate (lower) standard. Indeed, if maintenance is neglected, the loss sustained on a high standard road may be greater because the investment is larger. The proper solution is to ensure that maintenance is adequate to protect the investment made in a road to whatever standard it is designed.

11.2 Approach to selecting design standards

Most countries design their roads according to a 'design speed', which varies depending on the functional class of road and type of terrain. There is some evidence that the concept of design speed is not very appropriate as the basis for geometric design, as it will often lead to uneconomic design (Robinson 1981). Designs should be justified economically, and the optimum geometric standards will vary with both construction and road user costs. Construction costs will be related to terrain type and choice of pavement, while road user costs are directly related to the level and composition of traffic. For example, the higher the geometric design standards, the flatter must be the gradients on the road, which results in an increase in the earthworks construction costs, but a reduction in vehicles' fuel costs. The aim should be to design the gradients such that the marginal increase in earthworks costs is exactly

Stage 2:
providing
additional
capacity

Stage 3:
increase of
operational
efficiency
Application of
standards

Determination
of standards

Aims of the approach	<p>balanced by savings in vehicle operation costs. One criterion that can be used to determine geometric design standards is that of minimizing total cost; that is, the sum of construction, maintenance and road user costs.</p> <p>Even where national road design standards exist, it is recommended that the design process outlined in <i>Overseas Road Note 6</i> should be adopted to ensure that appropriate geometric design standards fulfil three inter-related objectives:</p> <ol style="list-style-type: none"> 1 to provide minimum levels of safety and comfort to road users by the provision of adequate sight distances, coefficients of friction and road space for vehicle manoeuvres; 2 to provide the framework for economic design; 3 to ensure a consistency of alignment.
Steps in the approach	<p>Furthermore, the design standards should take into account the environment of the road, traffic characteristics and driver behaviour.</p> <p>An outline of the process, which is intended to result in sound economic design, is shown in Figure 11.1. The design involves the following steps:</p> <ol style="list-style-type: none"> 1 Traffic flow, terrain type, and road function are defined, and a design class is chosen. 2 Trial alignments are selected, with design undertaken over sections with minimum lengths of about 1 km, by applying a series of discrete geometric elements of horizontal and vertical curvature. 3 Elements of lower geometric standard are identified and compared with the standards of the design class chosen. 4 Estimates of approach speeds are made for the geometric design elements identified above; if they are consistent, the design goes forward to economic evaluation; if not, the road alignment may be amended or the standards relaxed with the appropriate measures taken for safety.
Safety	<p>The design process recommended in <i>Overseas Road Note 6</i> aims to provide safe designs by ensuring consistency of design elements with speed, so that drivers do not face an unexpected change. The alignment design should include adequate stopping sight distances for the prevailing speeds and road surfaces, road space for vehicle manoeuvres, and clear signing and road marking. Segregation of motorized and non-motorized traffic is also recommended to improve safety.</p>

11.3 Classification of roads

11.3.1 Road classification

Geometric standards depend on the functional requirements of the road. The network is therefore divided into various classes of roads. The classification is also useful for road management purposes. A typical classification is shown in Box 11.1. While this classification of ‘access’, ‘collector’ and ‘arterial’ is simplistic, there will in practice be many overlaps of function, and clear distinctions will not always be apparent on functional terms alone. The classification should not be confused with the division of administrative responsibilities, which may be based on historical or political considerations.

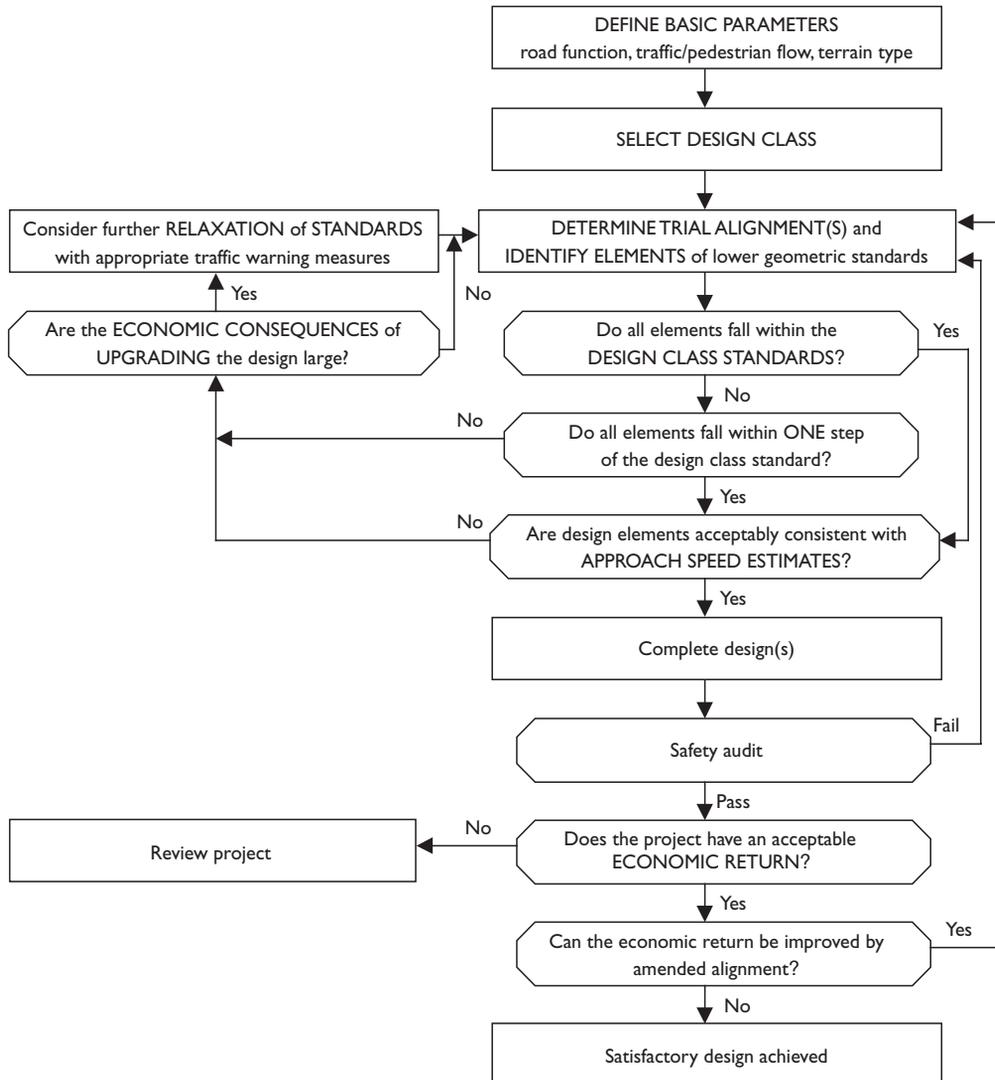


Figure 11.1 Design process.

Source: TRRL Overseas Unit (1988).

11.3.2 Design class

Overseas Road Note 6 recommends that the first step in the process should be to define the basic parameters of road function, terrain type and traffic flow. On the basis of these parameters, a design class is selected. Design speed is used only as an index which links design class to the design parameters of sight distance and curvature to ensure that a driver is presented with a reasonable consistent speed environment. Table 11.1 shows the design classes and design speeds recommended in *Overseas Road Note 6* in relation to road function, volume of traffic and terrain. The table also contains recommended standards for carriageway and shoulder width.

ORN6
approach

Box 11.1 Example of functional classification

Access roads

Access roads are the lowest level in the network hierarchy. Vehicular flows will be very light and will be aggregated in the collector road network. Geometric standards may be low and need only be sufficient to provide appropriate access to the rural agricultural, commercial and population centres served. Substantial proportions of the total movements are likely to be by non-motorized traffic.

Collector roads

Collector roads have the function of linking traffic to and from rural areas, either direct to adjacent urban centres, or to the arterial road network. Traffic flows and trip lengths will be of an intermediate level and the need for high geometric standards is therefore less important.

Arterial roads

Arterial roads are the main routes connecting national and international centres. Traffic on them is derived from that generated at the urban centres and from the inter-urban areas through the collector and access road systems. Trip lengths are likely to be relatively long and levels of traffic flow and speeds relatively high. Geometric standards need to be adequate to enable efficient traffic operation under these conditions, in which vehicle-to-vehicle interactions may be high.

Source: TRRL Overseas Unit (1988).

Table 11.1 Road design standards

Road function	Design class	Traffic flow ^a (ADT)	Surface type	Width (m)		Max gradient (%)	Terrain/design speed (km/h)		
				Carriage way	Shoulder		Mount	Rolling	Level
Arterial	A	5,000–15,000	Paved	6.5	2.5	8	85	100	120
	B	1,000–5,000	Paved	6.5	1.0	8	70	85	100
Collector	C	400–1,000	Paved	5.5	1.0	10	60	70	85
	D	100–400	Paved/unpaved	5.0	1.0 ^b	10	50	60	70
Access	E	20–100	Paved/unpaved	3.0	1.5 ^b	15	40	50	60
	F	< 20	Paved/unpaved	2.5/3.0	Passing places	15/20	N/A	N/A	N/A

Adapted from: TRRL Overseas Unit (1988).

Notes

- a The two-way traffic flow is recommended to be not more than one design class step in excess of first year ADT.
- b For unpaved roads where the carriageway is gravelled, the shoulders would not normally be gravelled; however, for Design Class D roads, considerations should be given to graveling the shoulders if shoulder damage occurs.

Table 11.2 Design speed and design class

Type	Number of lanes	Formation width (m)	Design speed (km/h)			Class			
			Flat	Rolling	Mount	A	B	C	D
I	4	36.0	120	100	80	A			
II	4	23.8	110	90	70	A	B		
III	2	14.4	100	90	60	A	B	C	
IV	2	12.4	100	80	60	A	B	C	
V	2	11.1	90	80	50	A ^a	B	C	
VI	2	9.7	90	80	50	A ^a	B ^a	C	D ^b
VII	2	9.0	80	70	50		B ^a	C	D ^b
VIII	2	7.5	70	70	50			C ^a	D
IX	1	7.0	70	60	40			C ^a	D

Source: SATCC Technical Unit (1995).

Notes

A = principal arterial roads; B = minor arterial roads; C = collector roads; D = access roads.

a Classification valid in SATCC countries.

b If not paved.

Terrain can be classified as ‘level’, ‘rolling’ or ‘mountainous’ based on the average ground slope measured as the number of 5-m contour lines crossed per kilometre on a straight line linking the two ends of the road section as follows:

Terrain

- Level terrain: 0–10 ground contours per kilometre
- Rolling terrain: 11–25 ground contours per kilometre
- Mountainous terrain: > 25 ground contours per kilometre.

Table 11.2 shows the design speed and design class according to the recommendations for the SATCC countries.

SATCC approach

11.4 Sight distance

11.4.1 Visibility requirements

The driver’s ability to see ahead contributes to safe and efficient operation of the road. Ideally, geometric design should ensure that, at all times, any object on the pavement surface is visible to the driver within normal eyesight distance. However, this is not usually feasible because of topographical and other constraints, so it is necessary to design roads on the basis of lower, but safe, sight distances. There are four different sight distances, which are of interest in geometric design:

- Stopping sight distance
- Meeting sight distance
- Passing sight distance
- Intersection sight distance.

11.4.2 Stopping sight distance

Components Stopping sight distance comprises two elements:

- d_1 = the distance moved from the instant the object is sighted to the moment the brakes are applied (the perception and brake reaction time, referred to as the total reaction time);
- d_2 = the distance traversed while braking (the braking distance).

Reaction time The total reaction time depends on the physical and mental characteristics of the driver, atmospheric visibility, types and condition of the road, distance to the hazard, plus its size, colour and shape. When drivers are keenly attentive, as in urban conditions with high traffic intensity, the reaction time may be in the range of 0.5–1.0 sec. Driver reaction time is generally around 2–4 sec for normal driving at rural conditions. *Overseas Road Note 6* assumes a total reaction time of 2 sec, while the SATCC recommendations assume 2.5 sec. The distance travelled before the brakes are applied, d_1 , is

$$d_1 = \frac{10}{36} \times V \times t$$

where d_1 is the total reaction distance in metres; V the initial vehicle speed in km/h; and t the reaction time in seconds.

Braking distance The braking distance, d_2 , is dependent on vehicle condition and characteristics, the coefficient of friction between tyre and road surface, the gradient of the road and the initial vehicle speed.

$$d_2 = \frac{V^2}{254 \times (f + g/100)}$$

where d_2 is the braking distance in metres; V the initial vehicle speed in km/h; f the coefficient of longitudinal friction; and g the gradient in per cent (positive if uphill and negative if downhill).

Longitudinal friction The determination of design values of longitudinal friction, f , is complicated because of the many factors involved. It is, however, known that f -values decrease for higher vehicle speeds on wet roads. The design values for longitudinal friction used in *Overseas Road Note 6* and in the SATCC Recommendations are shown in Table 11.3.

Recommendations The calculated stopping sight distances ($d_1 + d_2$) are listed in Tables 11.4 and 11.5.

11.4.3 Meeting sight distance

Meeting sight distance is the distance required to enable the drivers of two vehicles travelling in opposite directions to bring their vehicles to a safe stop after seeing each

Table 11.3 Coefficient of longitudinal friction, f

Design speed (km/h)	TRRL Overseas Road Note 6	SATCC recommendations
30	0.60	—
40	0.55	0.40
50	0.50	—
60	0.47	0.38
70	0.43	—
80	—	0.36
85	0.40	—
100	0.37	0.34
120	0.35	0.32

Adapted from: TRRL Overseas Unit (1988) and SATCC Technical Unit (1995).

Table 11.4 Stopping sight distance on level grade

Design speed (km/h)	Stopping sight distance (m)
30	25
40	35
50	50
60	65
70	85
85	120
100	160
120	230

Source: TRRL Overseas Unit (1988).

Table 11.5 Stopping sight distance for different gradients

Design speed (km/h)	Assumed speed for upgrades (km/h)	Coefficient of friction	Minimum stopping sight distance (m)				
			Upgrades		Level grade	Downgrades	
			+6%	+3%		−3%	−6%
40	35	0.40	44	44	44	44	47
60	55	0.38	74	76	79	83	86
80	70	0.36	116	122	126	133	140
100	85	0.34	173	178	185	197	210
120	100	0.32	242	250	260	279	301

Source: SATCC Technical Unit (1995).

other. Meeting sight distance is normally calculated as twice the minimum stopping sight distance. It is desirable that meeting sight distance is achieved throughout the length of single lane roads. It should also be provided on roads less than 5.0 m wide, since it is necessary to reduce speed to pass safely on such narrow roads.

11.4.4 Passing sight distance

Factors	Factors affecting passing (overtaking) sight distance are the judgement of overtaking drivers, the speed and size of overtaken vehicles, the acceleration capabilities of overtaking vehicles and the speed of oncoming vehicles. Driver judgement and behaviour are important factors that vary considerably among drivers. For design purposes, the passing sight distance selected should be adequate for the majority of drivers.
Determination	Passing sight distances are determined empirically and are usually based on passenger car requirements. Heavy commercial vehicles require longer time than cars to complete the overtaking manoeuvre but, on the other hand, commercial vehicle drivers can see further ahead because of their higher eye height when driving. Hence, they are able to judge sooner and better whether or not a gap is suitable for overtaking. This partially offsets any additional overtaking length that might be required.
Examples	There are considerable differences in various standards for passing sight distance due to different assumptions about the component distances in which a passing manoeuvre can be divided, different assumed speeds for the manoeuvre and, to some extent, driver behaviour. The passing sight distances recommended by <i>Overseas Road Note 6</i> and the SATCC recommendations are shown in Table 11.6. The passing sight distances recommended for use by the SATCC countries are based on a speed difference of 20 km/h between the passing vehicle and the overtaken vehicle, and acceleration rates proposed by AASHTO (1990). The reduced passing sight distance is based on the assumption that the overtaking vehicle may safely abandon the manoeuvre if an approaching vehicle comes into view.

11.4.5 Intersection sight distance

Intersection sight distance is the distance along the main road at which an approaching vehicle must be seen to permit a vehicle on an intersecting road to cross or merge safely with the traffic on the main road. The intersection sight distance depends on the design speed, the width of road being crossed and the characteristics of the

Table 11.6 Passing sight distances

Design speed (km/h)	TRRL Overseas Road Note 6	SATCC recommendations	
		Normal	Reduced
40	Not applicable	110	90
50	140	—	—
60	180	230	180
70	240	—	—
80	—	420	310
85	320	—	—
100	430	700	520
120	590	1,040	770

Adapted from: TRRL Overseas Unit (1988) and SATCC Technical Unit (1995).

vehicle crossing or merging on to the main road. Generally, the minimum intersection sight distance can be taken as:

$$S = 2 \times V, \text{ for passenger cars}$$

$$S = 3 \times V, \text{ for trucks}$$

where S is the sight distance along main road in metres and V the design speed in km/h.

11.5 Traffic

Information on traffic volumes, traffic composition and traffic loading is important in the determination of the appropriate standard of a road. The traffic has a major impact on the selection of road class, and consequently on all geometric design elements. Traffic types are described in Chapter 3 of this book, but the following are important for geometric design purposes:

- *Average annual daily traffic (AADT)* tends to be used for low volume roads, with the design control being the ‘design year’; for routes with large seasonal variations the design control is the average daily traffic (ADT) during the peak months of the ‘design year’ – the design year is usually selected to be ten years after the year of opening to traffic.
- *30th-hour volume* It is not economic to design a facility to be congestion-free at all times throughout the year; it is considered good practice to design roads to carry the ‘30th-hour volume (30th HV) – this is the hourly volume exceeded by only 29 hours per year.
- *Design hourly volume (DHV)* is expressed as $DHV = AADT \times K$ or $DHV = ADT \times K$, where K is estimated from the ratio of the 30th HV to the AADT from a similar site.

11.6 Cross-section

11.6.1 General principles

Road width should be minimized to reduce the costs of construction and maintenance, while being sufficient to carry the traffic loading efficiently and safely. The following factors need to be taken into account when selecting the width of a road:

Road width

- *Classification of the road* – a road is normally classified according to its function in the road network; the higher the class of road, the higher the level of service expected and the wider the road will need to be.
- *Traffic* – heavy traffic volumes on a road means that passing of oncoming vehicles and overtaking of slower vehicles are more frequent, and therefore the paths of vehicles will be further from the centre-line of the road, so the traffic lanes should be wider.
- *Vehicle dimensions* – normal steering deviations and tracking errors, particularly of heavy vehicles, reduce clearances between passing vehicles; wider traffic lanes are needed when the proportion of trucks is high.

- *Vehicle speed* – as speeds increase, drivers have less control of the lateral position of vehicles, reducing clearances, and so wider traffic lanes are needed.

Typical
recommen-
dations

Figure 11.2 shows the typical cross-sections recommended by *Overseas Road Note 6*, for different road design classes. These design classes and related traffic volumes were shown in Table 11.1.

11.6.2 Carriageways

Single-lane
access roads

For access roads with low volumes of traffic (AADT < 20), single-lane operation is adequate as there will be only a small probability of vehicles meeting. The few passing manoeuvres can be undertaken at very reduced speeds using either passing places or shoulders. Provided that meeting sight distance is available, these manoeuvres can be performed without hazard, and the overall loss in efficiency brought about by the reduced speeds will be small as only a few such manoeuvres will be involved.

Two-lane
access roads

For higher traffic flows (20–100 vehicles per day), single track roads cause considerable inconvenience to traffic and are only recommended for short roads, or in hilly or mountainous terrain. In these cases, there are high costs of construction in side cut and for haulage of materials. Most countries with a rural access road programme have provided roads that are sufficiently wide for two vehicles to pass safely. It should also be noted that, if the new road is to be constructed by machine, the extra construction cost of building a wider road would be quite small in level and rolling terrain. On unpaved roads, it is not normally possible to distinguish between the running surface and the shoulder. Where the formation material has insufficient bearing capacity, the surface should be gravelled to the top of the side slopes.

Example

The Public Works Department in Bhutan, a very mountainous country, recommends single-lane pavements of 3.0–3.5 m width for traffic volumes up to 200 vehicles per day, and with shoulder widths of 0.5–2.0 m, as shown in Table 11.7.

Collector
roads

On roads with medium volumes of traffic (100–1,000 vehicles per day), the numbers of passing manoeuvres will increase and pavement widening will become worthwhile operationally and economically. However, in view of the generally high cost of capital for construction in developing and emerging countries and the relatively low cost of travel time, reductions in speed of approaching vehicles is likely to be acceptable for such flow levels. Running surface widths of 5.0 and 5.5 m are recommended in *Overseas Road Note 6*. However, in case of high percentage of heavy vehicles (> 40 per cent), it is advisable to increase the running surface width to 5.5 or 6.0 m.

Arterial roads

For arterial roads with higher flows (> 1,000 vehicles per day), a 6.5 m wide running surface will allow heavy vehicles to pass safely without the need to move laterally in their lanes or to slow down. The standard cross-sections proposed in the SATCC recommendations are listed in Table 11.8.

Culverts and
bridges

The cross-section of the road is usually maintained across culverts, but special cross-sections may need to be designed for bridges, taking into account traffic such as pedestrians, cyclists, as well as motor traffic. Reduction in the carriageway width may be accepted when an existing narrow bridge has to be retained because widening or replacement may not be economically feasible. It may also sometimes be economic to construct a superstructure of reduced width initially with provision

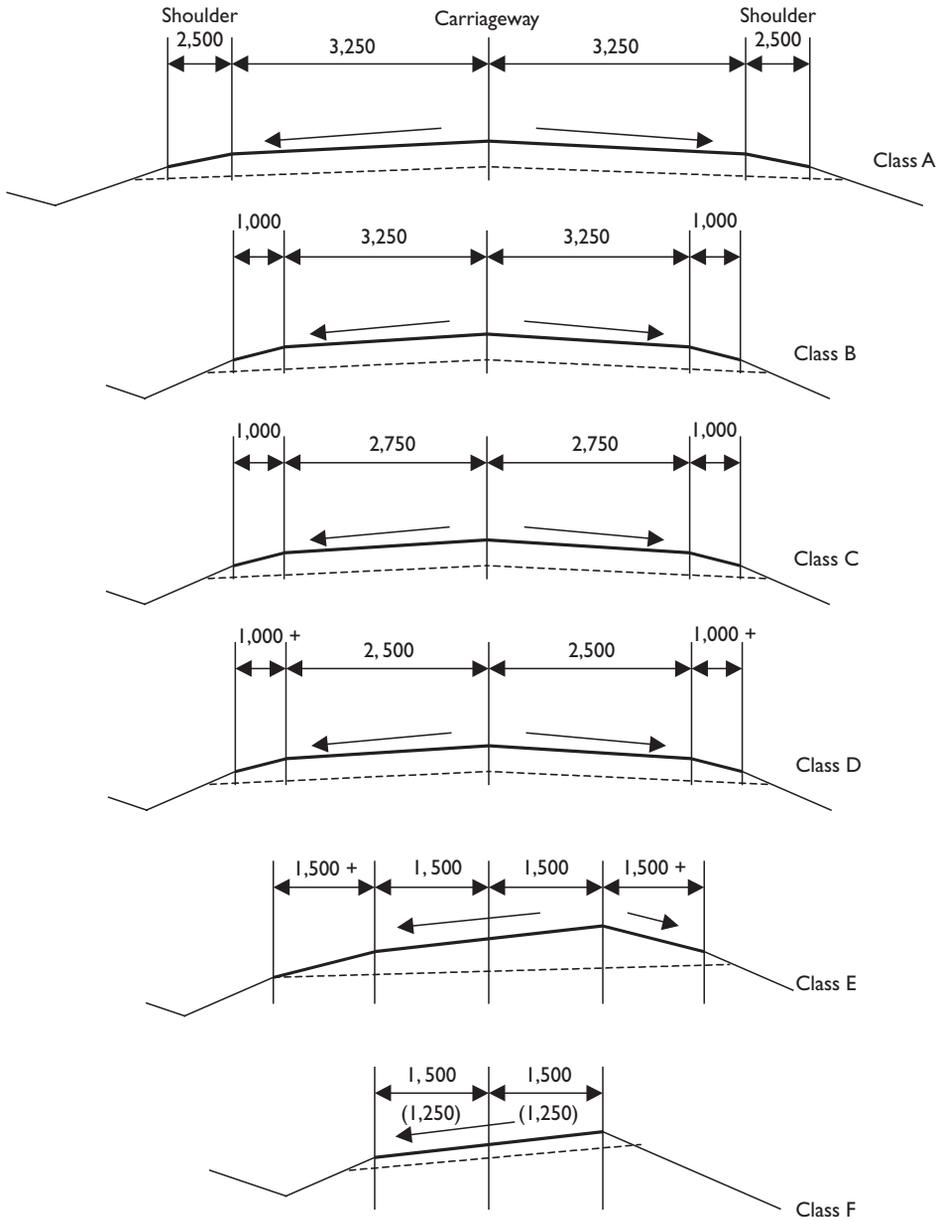


Figure 11.2 Typical cross-sections.

Source: TRRL Overseas Unit (1988).

Notes

- (1) Dimensions are in mm.
- (2) Widths should be considered as minimum values and widening is required on curves with tighter radii.
- (3) Single-lane road (Classes E and F) require meeting sight distance.
- (4) For high percentage of heavy vehicles (> 40 per cent), it is advisable to increase running surface width for class C, D, E and F by 0.50m.

Table 11.7 Pavement, shoulder and formation widths of rural roads in Bhutan

Road class (design traffic volume)		Design speed (km/h)	
		Up to 60	Over 60
Class AA Two lane (> 200)	Pavement	6.0	6.5
	Shoulder	1.0	1.5
	Formation	8.0	9.5
Class A One lane (100–200)	Pavement	3.5	3.5
	Shoulder	1.5	2.0
	Formation	6.5	7.5
Class B One lane (50–100)	Pavement	3.5	3.5
	Shoulder	1.0	1.5
	Formation	5.5	6.5
Class C One lane (< 50)	Pavement	3.0	3.0
	Shoulder	0.75	1.25
	Formation	4.5	5.5
Class D One lane (< 25)	Pavement	3.0	—
	Shoulder	0.5	—
	Formation	4.0	—

Source: Snowy Mountains Engineering Corporation (undated).

Table 11.8 Standard cross-sections for rural roads

Road type	No. of lanes	Lane width (m)	Edge strip (m)	Carriageway width (m)	Shoulder width (m)	Outer shoulder (m)	Median width (m)	Formation width (m)
I	4	3.50	0.5	2 × 8.0	2.5 ^a	1.5	12	36.0
II	4	3.50	0.2	2 × 7.4	2.0 ^a	1.0	3 ^b	23.8
III	2	3.50	0.2	7.4	2.5 ^a	1.0		14.4
IV	2	3.50	0.2	7.4	2.0 ^a	0.5		12.4
V	2	3.35	0.2	7.1	1.5	0.5		11.1
VI	2	3.35	—	6.7	1.5			9.7
VII	2	3.00	—	6.0	1.5			9.0
VIII	2	2.75	—	5.5	1.0			7.5
IX	1	3.00	—	3.0	2.0			7.0

Source: SATCC Technical Unit (1995).

Notes

a Emergency lane (2.5 or 2.0m wide) to be paved for principal arterial roads.

b Median with guard rail.

for it to be widened later when warranted by traffic. In such cases, a proper application of traffic signs, rumble strips or speed humps is required to warn motorists of the discontinuity in the road width.

Passing places

For single-lane roads without shoulders, passing places must be provided to allow passing and overtaking to occur. The total road width at passing places should be minimum 5.0m, and preferably 5.5m, which allows two trucks to pass safely at low speed.

Table 11.9 Curve widening (TRRL)

Single-lane roads with 3.0m basic width				
Curve radius (m)	20	30	40	60
Increase in width (m)	1.50	1.00	0.75	0.50
Two-lane roads				
Curve radius (m)	< 50	50–149	150–299	300–500
Increase in width (m)	1.50	1.00	0.75	0.50

Source: TRRL Overseas Unit (1988).

Table 11.10 Curve widening (SATCC)

Curve radius (m)	Lane widening in metres for normal lane width (m)			
	2.75	3.00	3.35	3.50
< 50	1.00	1.00	0.65	0.50
50–100	0.75	0.75	0.40	0.25
100–250	0.50	0.50	0.25	
250–750	0.25	0.25		

Source: SATCC Technical Unit (1995).

The length of individual passing places will vary with local conditions and the sizes of vehicles in common use but, generally, a length of 20m including tapers will cater for trucks with a wheel base of 6.5m and overall length of 11m. Normally, passing places should be located every 300–500m along the road, depending on the terrain and geometric condition. They should be located within view of each other, and be constructed at the most economic locations as determined by terrain and ground conditions, such as at transitions from cut to fill, rather than at precise intervals.

Carriageway widths should be increased on low radius curves to allow for the swept paths of longer vehicles, and the necessary tolerances in lateral location as vehicles follow a curved path. Widening should be applied on the inside of a curve and be introduced gradually over the length of transition curve (see Chapter 12). The amount of widening recommended by *Overseas Road Note 6* for single-lane and two-lane roads are shown in Table 11.9. The lane widening on curves (widening of each lane) according to the SATCC recommendations are shown in Table 11.10.

On access roads, which often have substantial horizontal curvature requiring local widening, it may be practical to increase the width over a complete section to offer a more consistent aspect to the driver. This enhancement of the standards should be undertaken where other advantages such as easier construction or maintenance can be identified and where the additional costs are acceptably small.

Curve widening

Access roads

11.6.3 Shoulders

Shoulders provide for the accommodation of stopped vehicles. Properly designed shoulders also provide an emergency outlet for motorists finding themselves on

Purpose

a collision course and they also serve to provide lateral support for the carriageway. Further, shoulders improve sight distances and induce a sense of ‘openness’, which improves capacity and encourages uniformity of speed. Shoulders are recommended on all classes of road, except minor access roads in hilly or mountainous terrain where they would increase cost considerably. In developing and emerging countries, shoulders are used extensively by non-motorized traffic (pedestrians, bicycles and animals) and a significant proportion of the goods may be transported by such non-motorized means.

Recommendations

For access roads, the combined width of carriageway and shoulders is recommended to be 6 m, which is sufficient for two trucks to pass with 1-m clearance. A minimum of one metre of surfaced shoulders is recommended for paved collector and arterial roads to avoid damage and the resulting break up of the edge of the carriageway pavement. For roads with high traffic flow, 2.5-m wide shoulders should be provided as stopped vehicles blocking any part of the carriageway would cause a significant hazard. Of the 2.5-m shoulder, at least 1 m should be paved. Paved shoulders should be clearly segregated from the carriageway either by use of 100-mm wide carriageway edge line marking, or by sealing the shoulders with a different coloured aggregate.

11.6.4 Crossfall

Principles for carriageways

Two-lane roads should be provided with a camber consisting of a straight line crossfall from the centre-line to the carriageway edges, while straight crossfall from edge to edge of the carriageway can be used for single-lane roads and for each carriageway of divided roads. The crossfall should be sufficient to provide adequate surface drainage whilst not being so great as to be hazardous by making steering difficult. *Overseas Road Note 6* recommends a normal crossfall of 3 per cent on paved roads and 4–6 per cent on unpaved roads. The ability of a surface to shed water varies with its smoothness and integrity. On unpaved roads, the minimum acceptable value of crossfall should be related to the need to carry surface water away from the pavement structure effectively, with a maximum value above which erosion of material starts to become a problem. The crossfall of unpaved roads will reduce over time due to action of traffic and weather, and rutting may develop. To avoid the ruts developing into pot-holes, a crossfall of 5–6 per cent should be re-established during recurrent and periodic maintenance works.

Shoulders

Unpaved shoulders on a paved road should be about 2 per cent steeper than the crossfall of the carriageway to provide effective drainage. For practical construction reasons, it may be desirable to apply the same cross slope on the shoulder as on the carriageway, if both carriageway and shoulder are paved with the same type of material.

11.6.5 Superelevation

Need

With normal crossfall, vehicles on the outside lane of a horizontal curve need to develop high levels of frictional force to resist sliding, depending on the speed, curve radius and crossfall. For small radius curves and at higher speeds, the removal of adverse crossfall alone will be insufficient to reduce frictional needs to

an acceptable level, and ‘superelevation’ should be applied. The preferable maximum value of superelevation is normally set at 7 per cent, with an absolute maximum of 10 per cent, to take account of the stability of slow, high laden commercial vehicles and the appearance of the road.

The development of superelevation for two-lane roads is normally achieved by raising the outside of the pavement edge relative to the centre-line at a constant relative longitudinal gradient until a straight cross-section is obtained. The whole cross-section is then rotated about the centre-line until the full superelevation is obtained. The relative longitudinal gradient, that is, the difference between the outside pavement edge and the centre-line, should not exceed a maximum value established from considerations of appearance and comfort. Superelevation design curves are shown in Figure 11.3. On paved roads with unsealed shoulders, the shoulders should drain away from the paved area to avoid loose material being washed across the road. Superelevation is normally provided so that two-thirds are applied prior to the start of the circular curve, and one-third inside the curve. However, where transition curves are used, superelevation should be applied over the length of the transition curves.

The transition length from a normal cross-section on a tangent to the fully super-elevated cross-section, called the ‘superelevation run-off’ length, is directly proportional to the total superelevation according to the relationship:

$$L = \frac{W}{2S}(e_0 + e)$$

where L is the superelevation run-off (m); W the width of carriageway (m); S the constant relative longitudinal gradient (per cent); e the superelevation of the curve (per cent); and e_0 the normal crossfall on the straight (per cent).

Typical maximum and minimum values for the relative longitudinal gradient, S , are given in Table 11.11.

Application

Superelevation run-off

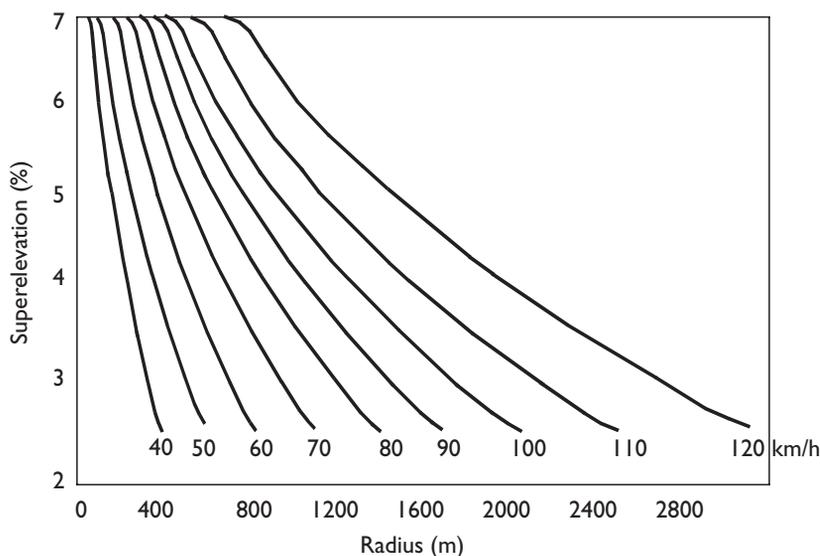


Figure 11.3 Superelevation design curves.

Source: TRRL Overseas Unit (1988).

Table 11.11 Maximum and minimum values for the relative longitudinal gradient

Design speed (km/h)	Relative longitudinal gradient (%)		
	Maximum (SATCC)	Maximum (Botswana)	Minimum
≤ 40	1.50	0.71	0.3
50	1.25	0.67	0.3
60	1.00	0.59	0.3
70	0.75	0.54	0.3
80	0.50	0.50	0.3
90	0.50	0.46	0.3
100	0.50	0.43	0.3

Adapted from: SATCC Technical Unit (1995), Botswana MoWC (1982).

Table 11.12 Suggested earth slopes for design

Element	Access roads	Arterial and collector roads
Inner slope of side drain	1:3	1:4
Low fills (height ≤ 3 m)	1:2	1:4
High fills (height > 3 m)	1:1.5	1:2
Low cuts (height ≤ 3 m)	1:1.5	1:2
High cuts (height > 3 m)	1:1	1:1.5

11.6.6 Side slopes

- Requirements** The slopes of embankments and cuttings must be adapted to the soil properties, topography and importance of the road. Earth fills of common soil types are usually stable at slopes of 1:1.5, providing that the embankments are not too high. Slopes of cuttings through ordinary undisturbed earth with cementing properties may be stable with slopes of about 1:1. Rock cuts are usually stable at slopes of 4:1, or even steeper, depending on the homogeneity of the rock formation and direction of possible dips and strikes.
- Erosion** Using these relatively steep slopes will minimize earthworks, but they are more liable to erosion than flatter slopes, since vegetation growth is hampered and surface water velocity will be higher. Thus, the savings in original excavation and embankment costs may be more than offset by increased maintenance over the life of the road. If construction is to be undertaken by labour-based methods, earthworks should be minimized by the use of relatively steep slopes consistent with the angle of repose of the material.
- Accident hazards** Steep slopes on fills and inner slopes of side drains can create accident hazards. Drivers can lose control and overturn if the wheel of a vehicle strays over the edge of the shoulder. With flatter slopes, the vehicle can be steered back onto the road, or continue down the side slope without overturning.
- Recommendations** Suggested cut and fill slopes in earth which generally yield favourable cross-sections are given in Table 11.12.

References

- AASHTO (1984, 1990) *A Policy on Geometric Design of Highways and Streets*, Washington, DC: American Association of State Highway and Transportation Officials.
- Botswana MoWC (1982) *Botswana Road Design Manual*, Gaborone: Ministry of Works and Communications, Botswana.
- Boyce, A.M., McDonald, M., Pearce, M.J. and Robinson, R. (1988) 'A review of geometric design and standards for rural roads in developing countries', *Contractor Report 94*, Crowthorne: Transport and Road Research Laboratory.
- Department of Transport (1981) 'Highway link design', *Departmental Standard 9/81*. London: H.M. Stationery Office.
- Kosasih, D., Robinson, R. and Snell, J. (1987) 'A review of some recent geometric road standards and their application to developing countries', *TRRL Research Report 114*, Crowthorne: Transport and Road Research Laboratory.
- McLean, J.R. (1978) 'The role of geometric standards', in REAAA *Second Conference of REAAA, Manila, 16–20 October 1978, Volume 2 (Design standards)*, no place of publication cited: Road Engineering Association of Asia and Australasia, 65–75.
- NAASRA (1980) *Interim Guide to the Geometric Design of Rural Roads*, Sydney: National Association of Australian State Road Authorities.
- Robinson, R. (1981) 'The selection of geometric design standards for rural roads in developing countries' *TRRL Supplementary Report 670*, Crowthorne: Transport and Road Research Laboratory.
- SATCC Technical Unit (1995) *Recommendations on Road Design Standards, Volume I: Geometric Design of Rural Roads*, Maputo: Southern Africa Transport and Communications Commission.
- Snowy Mountains Engineering Corporation (undated) *Road Design Manual*, Thimphu: Public Works Department, Royal Government of Bhutan.
- TRRL Overseas Unit (1988) 'A guide to geometric design', *Overseas Road Note 6*, Crowthorne: Transport and Road Research Laboratory.

Geometric alignment design

Kent Falck-Jensen

12.1 Geometric design elements

Alignment and cross-section elements

The following elements must be considered when carrying out the geometric design of a road:

- Horizontal alignment:
 - minimum curve radius (maximum degree of curvature)
 - minimum length of tangent between compound or reverse curves
 - transition curve parameters
 - minimum passing sight distance and stopping sight distance on horizontal curves
- Vertical alignment:
 - maximum gradient
 - length of maximum gradient
 - minimum passing sight distance or stopping sight distance on summit (crest) curves
 - length of sag curves
- Cross-section:
 - width of carriageway
 - crossfall of carriageway
 - rate of superelevation
 - widening of horizontal curves
 - width of shoulder
 - crossfall of shoulder
 - width of structures
 - width of right-of-way
 - sight distance
 - cut and fill slopes and ditch cross-section.

The standards to be chosen for these design elements are dependent on the criteria for the geometric design controls, that is, design class, sight distance, traffic volume and level of service, as described in the previous chapter, including appropriate standards for cross-section.

The choice of standard for geometric design elements needs to be balanced to avoid some elements having minimum values when others are considerably above the minimum requirements. Balanced design

12.2 Horizontal alignment

12.2.1 Basis of design

The horizontal alignment consists of a series of intersecting tangents and circular curves, with or without transition curves. The horizontal alignment should always be designed to an economic standard consistent with the topography and be chosen carefully to provide good drainage and to minimize earthworks. The alignment design should also be aimed at achieving a uniform operating speed. Therefore, the standard of alignment selected for a particular section of road should extend throughout the section with no sudden changes from easy to sharp curvature. Where sharper curvature is unavoidable, a sequence of curves of decreasing radius is recommended. In particular, near-minimum curves should not be used at the following locations:

- On high fills or elevated structures as the lack of surrounding features reduces the drivers perception of the alignment.
- At or near vertical curves, especially crest curves, as the unexpected bend can be extremely dangerous, particularly at night time.
- At the end of long tangents or a series of gentle curves; compound curves, where a sharp curve follows a long flat curve, should also be avoided to avoid misleading drivers.
- At or near intersections and approaches to bridges, in particular approaches to single lane bridges.

12.2.2 Tangents

Long straights should be avoided as they are monotonous and cause headlight dazzle. An improved and safer alternative is obtained by a winding alignment with tangents deflecting some 5–10 degrees alternately to the left and right. Short straights between curves in the same direction give a ‘broken-back’ appearance, and should not be used. Where reasonable tangent length is not attainable, the use of long transitions or compound curvature should be considered. The broken-back effect may also be avoided by the introduction of a sag curve. The following guidelines may be applied:

- Straights should have lengths less than $20 \times V$ m, where V is the design speed in km/h.
- Straights between circular curves turning in the same direction should have lengths greater than $6 \times V$ m, where V is the design speed in km/h.
- Straights between the end and the beginning of reverse circular curves, with no transition curve, should have lengths greater than two-thirds of the minimum of the total superelevation run-off (see Chapter 11).

12.2.3 Circular curves

When vehicles negotiate a circular curve, a sideways frictional force is developed between the tyres and road surface. This inertial force must be balanced by centripetal forces derived from the applied superelevation. The relationship between the radius, speed and frictional forces required to keep the vehicle in its path are given by:

$$R = \frac{V^2}{127(100e + f_s)}$$

where R is the radius of curve (m); V the speed of vehicle (km/h); e the crossfall of road (per cent) (e is negative for adverse crossfall); and f_s the coefficient of side (radial) friction force developed between the tyres and road pavement.

When vehicles brake while traversing a curve, both side and tangential frictional forces become active. The portion of the side friction factor that can be used with comfort and safety is normally determined as not more than half of the tangential coefficient of friction. Superelevation was discussed in Chapter 11.

12.2.4 Minimum radius

Table 12.1 shows the minimum horizontal curve radii together with assumed side friction factors recommended by different design guides. The recommended minimum radii of curves below which adverse crossfall should be removed is shown in Table 12.2. The minimum radius for design of horizontal curves used in Bhutan, which is also shown in Table 12.1, is based on superelevation and side friction

Table 12.1 Minimum radii for horizontal curves in metres

Design speed (km/h)	TRRL Overseas Road Note 6 ^a			SATCC Recommendations ^b		Bhutan ^c
	Side friction	Radius (e = 10%)	Radius (e = 7%)	Side friction	Radius (e = 7%)	Radius (e = 7%)
20	0.33	—	—	—	—	15
30	0.30	15	20	—	—	25
40	0.25	30	35	0.20	50	45
50	0.23	60	65	—	80	75
60	0.20	85	95	0.16	125	120
70	—	130	145	—	180	200
80	0.18	—	—	0.13	250	280
85	—	210	230	—	—	—
90	0.15	—	—	—	335	—
100	0.15	320	360	0.11	440	—
110	—	—	—	—	560	—
120	—	450	515	0.09	710	—

Adapted from

a TRRL Overseas Unit (1988).

b SATCC Technical Unit (1995).

c Snowy Mountains Engineering Corporation (undated).

Table 12.2 Minimum radii of curve with adverse crossfall in metres

Design speed (km/h)	TRRL Overseas Road Note 6 (e = -3%)	SATCC Recommendations (e = -2.5%)
< 50	500	1,800
60	700	1,800
70	1,000	2,500
80	—	3,000
85	1,400	—
90	—	4,000
100	2,000	5,000
110	—	6,000
120	2,800	7,000

Adapted from: TRRL Overseas Unit (1988), SATCC Technical Unit (1995).

considerations, as well as the need to ensure stopping sight distance in the curve. The maximum design values of side friction coefficient used in Bhutan vary from 0.19 at 20–40 km/h down to 0.12 at 80 km/h.

12.2.5 Visibility distance

Situations frequently exist where something on the inside of a curve, such as vegetation, a building or the face of a cutting, obstructs the line of sight. Where it is either not feasible or economically justified to move the object, a larger radius curve will be required to ensure that stopping sight distance is available. The required radius of curve is dependent on the distance of the obstruction from the centre-line and the sight distance, as shown in Figure 12.1. It can be derived from the relationship:

$$M = \frac{S^2}{8(R - N)} + N$$

where M is the obstruction to centre-line distance (m); R the radius of horizontal curve (m); S the stopping sight distance (m); and N the drivers eye and road object displacement from centre-line in metres (can be assumed to be 1.8 m).

12.2.6 Curve length

The horizontal alignment design should provide the maximum length of road with adequate sight distances and where overtaking can be carried out. This requires using curve radii close to the desirable minimum value to minimize the length of the horizontal curves. Earlier design methods used longer curves to produce ‘flowing alignments’ and more gentle bends. However, with such designs, there are restricted sight distances on the longer curves and a shorter length of alignment will be available where overtaking is safe.

However, for small changes of direction, it is often desirable to use large radius curves. This improves the appearance of the road by removing rapid changes in edge profile. It also reduces the tendency for drivers to cut the corners. The use of long

Design
philosophy

Other
situations

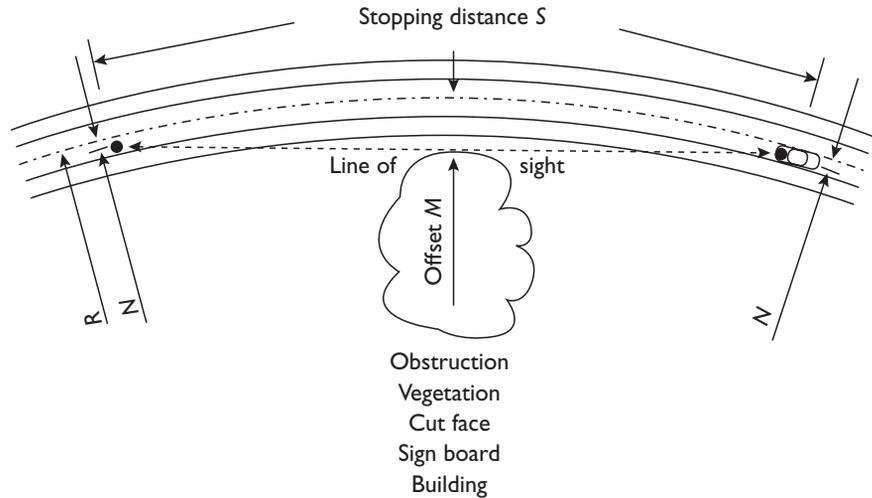


Figure 12.1 Minimum offset to obstruction adjacent to curve.

curves with radius near absolute minimum should be avoided. On these, drivers at speeds other than the design speed find it difficult to remain in lane. Curve widening can reduce such problems.

12.2.7 Transition curves

Purpose

The provision of transition curves between tangents and circular curves has the following benefits:

- Transition curves provide a natural easy-to-follow path for drivers, such that the centripetal force increases and decreases gradually as a vehicle enters and leaves a circular curve.
- The transition between the normal cross-slope and the fully superelevated section on the curve can be effected along the length of transition curve in a manner closely fitting the speed–radius relationship for the vehicle.
- Where the pavement is to be widened around a sharp circular curve, the widening can conveniently be applied over the transition curve length.
- The appearance of the road is enhanced by the application of transition curves.

Type of curve

Transition curves should be used on all superelevated curves for arterial (primary) roads with high design speed. The 'euler spiral', or 'clothoid', which is characterized by having a constantly changing radius, is normally considered the most convenient, and is defined by the equation:

$$A^2 = R \times L$$

where R is the circular curve radius in metres; L the length of clothoid in metres; and A the clothoid parameter in metres.

For aesthetic reasons, the clothoid should have a deflection angle of at least three degrees; thus:

$$A \geq R/3$$

Furthermore, the clothoid should be of sufficient length to accommodate the super-elevation run-off, thus:

$$A \geq \sqrt{R \times L_0} = \sqrt{R \times \frac{W}{2} \times \frac{e_1 - e_0}{S_{\max}}}$$

where R is the radius of circular curve (m); L_0 the minimum length of superelevation run-off (m); W the carriageway width (m); e_1 the superelevation of circular curve (per cent); e_0 the crossfall on the tangents (per cent); and S_{\max} the maximum rate of change of superelevation (per cent) – see Chapter 11.

12.3 Vertical alignment

12.3.1 Design elements

The vertical alignment of a road has a strong influence upon the construction cost, the operating cost of vehicles using the road, and in combination with the horizontal alignment also on the number of accidents. The vertical alignment should be designed to an economic standard consistent with the needs of traffic and the topography. Preferably, it should also be designed to be pleasing aesthetically. The two major elements of vertical alignment are the gradient, which is related to vehicle performance and level of service, and the vertical curvature, which is governed by sight distance and comfort criteria.

12.3.2 Gradient

Gradients need to be considered from the standpoint of both length and steepness, and the speed at which heavy vehicles enter the gradient. They should be chosen such that any marginal increase in construction cost is more than offset by the savings in operating costs of the heavy vehicles ascending them over the project analysis period. For access roads with low levels of traffic (less than about 20 vehicles per day), it is appropriate to use the maximum gradient that the anticipated type of vehicle can climb safely.

Maximum gradient is limited ultimately by traction requirements. Four-wheel drive vehicles can climb gradients in excess of 20 per cent, while small commercial vehicles can usually negotiate an 18 per cent gradient. Two-wheel drive trucks can tackle gradients of 15–16 per cent, except when heavily laden. These performance considerations provide the limiting criteria for the gradients shown in Table 12.3. These maximum traversable gradients are extremely steep and, if possible, all gradients should be much less severe.

On roads with an earth surface, particular soil types may give rise to slippery conditions during rains and even moderate gradients of around 5 per cent can be very

Basic
requirements

Absolute
maximum
gradients

Unpaved
roads

Table 12.3 Maximum recommended gradient

Road function	Design class	Maximum gradient (%)
Arterial	A	8
	B	8
	C	10
Collector	C	10
	D	10
Access	D	10
	E	15
	F	15/20

Source: TRRL Overseas Unit (1988).

Table 12.4 Maximum gradients depending on type of terrain

Design speed (km/h)	Maximum gradient (%)		
	Flat	Rolling	Mountainous
40	—	—	10
50	—	7	9
60	—	6	8
70	5	5.5	7
80	4	5	6
90	3.5	4.5	—
100	3	4	—
110	3	—	—
120	3	—	—

Source: SATCC Technical Unit (1995).

difficult to negotiate. On tight horizontal curves, and in particular on hairpin bends, it is desirable to use gradients well below the maximum values given in Table 12.3. Gradients of 10 per cent or over will usually need to be paved to enable sufficient traction to be achieved and to facilitate pavement maintenance.

Arterial roads

As traffic flows increase, the economic benefits of reduced vehicle operating and travel time may justify reducing the severity and/or length of gradients. The maximum gradients recommended by *Overseas Road Note 6* reflect the economics, as well as the need to avoid the build up of local congestion. However, separate economic assessment of alternatives to long or severe gradients should be undertaken, where possible and necessary, using models such as *HDM-4* (see Chapter 21).

Terrain

The cost of construction of a road generally increases as the terrain gets steeper, so the use of higher gradients is justified. This is reflected in the SATCC recommendations in Table 12.4.

Truck speeds

Sustained grades steeper than about 3 per cent slow down heavy vehicles significantly. Gradient design should aim at grades that will not reduce the speeds of heavy

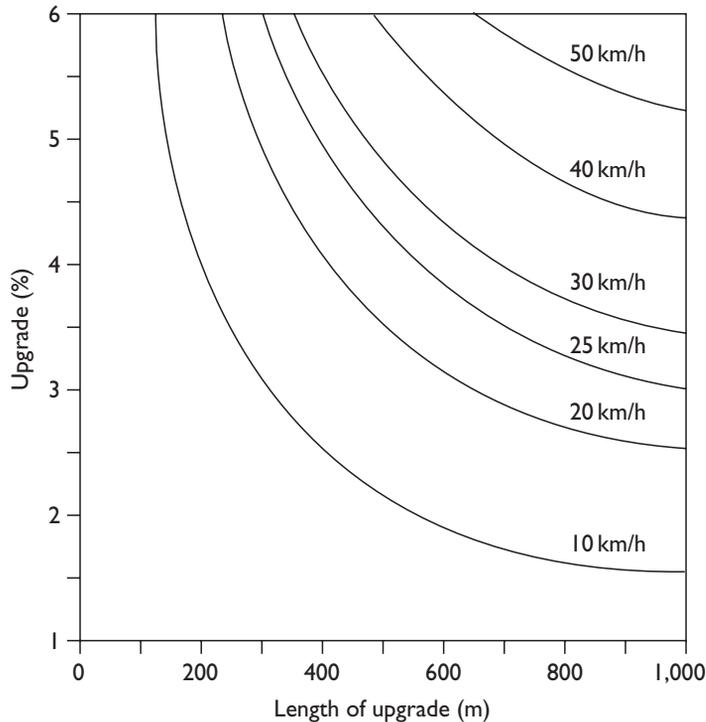


Figure 12.2 Speed reduction effect of upgrade.

Source: Botswana MoWC (1982).

vehicles by more than 25 km/h. Critical lengths of constant grade for various reductions in truck operating speed are given in Figure 12.2, assuming that the truck is entering the gradient at 80 km/h.

Climbing lanes enable faster vehicles to overtake more easily, resulting in shorter average journey times and reduced vehicle-operating costs. Benefits increase with increases in gradient, length of gradient, traffic volume, the proportion of trucks, and reductions in overtaking opportunities. Their use may be a cost-effective alternative where terrain or other physical features prevent shortening or flattening of gradients, and where the length of critical grade is exceeded.

Climbing lane

12.3.3 Vertical curves

Vertical curves are used to provide smooth transitions between consecutive gradients. A simple parabola is normally used to provide a constant rate of change of curvature, and hence visibility, along its length. It has the form

Parabolic curves

$$y = \frac{x^2}{2R}$$

where y is the vertical distance from the tangent to the curve (m); x the horizontal distance from start of vertical curve (m); and R the radius of the circular curve that is approximated by the parabola (m). x is assumed to be small relative to R .

For practical values of gradient, the length of parabolic curve, L_p , can be assumed to be equal to the chord length, L_C , and twice the tangent length, L_T . It may be approximated by the equation:

$$L_p = L_C = 2L_T = \frac{A}{100} R = \frac{i_1 - i_2}{100} R$$

where A is the algebraic difference between the gradients i_1 and i_2 of the intersecting tangents (per cent).

K-value

The rate of change of gradient along the curve is constant, and equals the algebraic difference in gradients (A), divided by the length of the curve in metres (L_C). The reciprocal, L_C/A , termed the 'K-value', is the horizontal distance in metres required to effect a one per cent change in gradient. It is, therefore, a measure of curvature. The K-value ($= L_C/A = R/100$) can be used to determine the distance from the start of the vertical curve to the apex of crest curves or the bottom of sag curves. This point where the gradient is zero occurs at a distance from the start of the vertical curve equal to K times the approach gradient.

12.3.4 Crest curves

Consistency of alignment

The driver's view of the road ahead is obscured by crest vertical curves, and it is not possible for them to assess the severity of the curvature further along the road. Drivers place a large measure of trust in the designer and assume that there will not be an unexpected tight curve that is inconsistent with the remainder of the alignment.

Sight distance

The minimum required lengths of crest curves are normally designed to provide stopping sight distance during daylight conditions. Longer lengths would be needed to meet the same visibility requirements at night on unlit roads. Even on a level road, low meeting-beam headlight illumination may not show up small objects at the design stopping sight distances. However, vehicle tail lights and taller objects will be illuminated at the required stopping sight distances on crest curves. Drivers are likely to be more alert at night and/or travelling at reduced speed, so longer lengths of curve are not justified.

Curve length

Required lengths of crest curves for safe stopping are shown in Figure 12.3. If full overtaking sight distance cannot easily be obtained, the design should aim to reduce the length of crest curves to provide the minimum stopping sight distance, thus increasing overtaking opportunities on the gradients on either side of the curve.

12.3.5 Sag curves

Visibility requirements

The curvature of sag curves is governed by visibility at night. This is limited by the distance illuminated by headlamp beams. The minimum curve length, L , for this condition, with provision of safe stopping sight distance, is given by the relationships:

For $S < L$:

$$L = \frac{A \times S^2}{200(h + S \times \tan \alpha)}$$

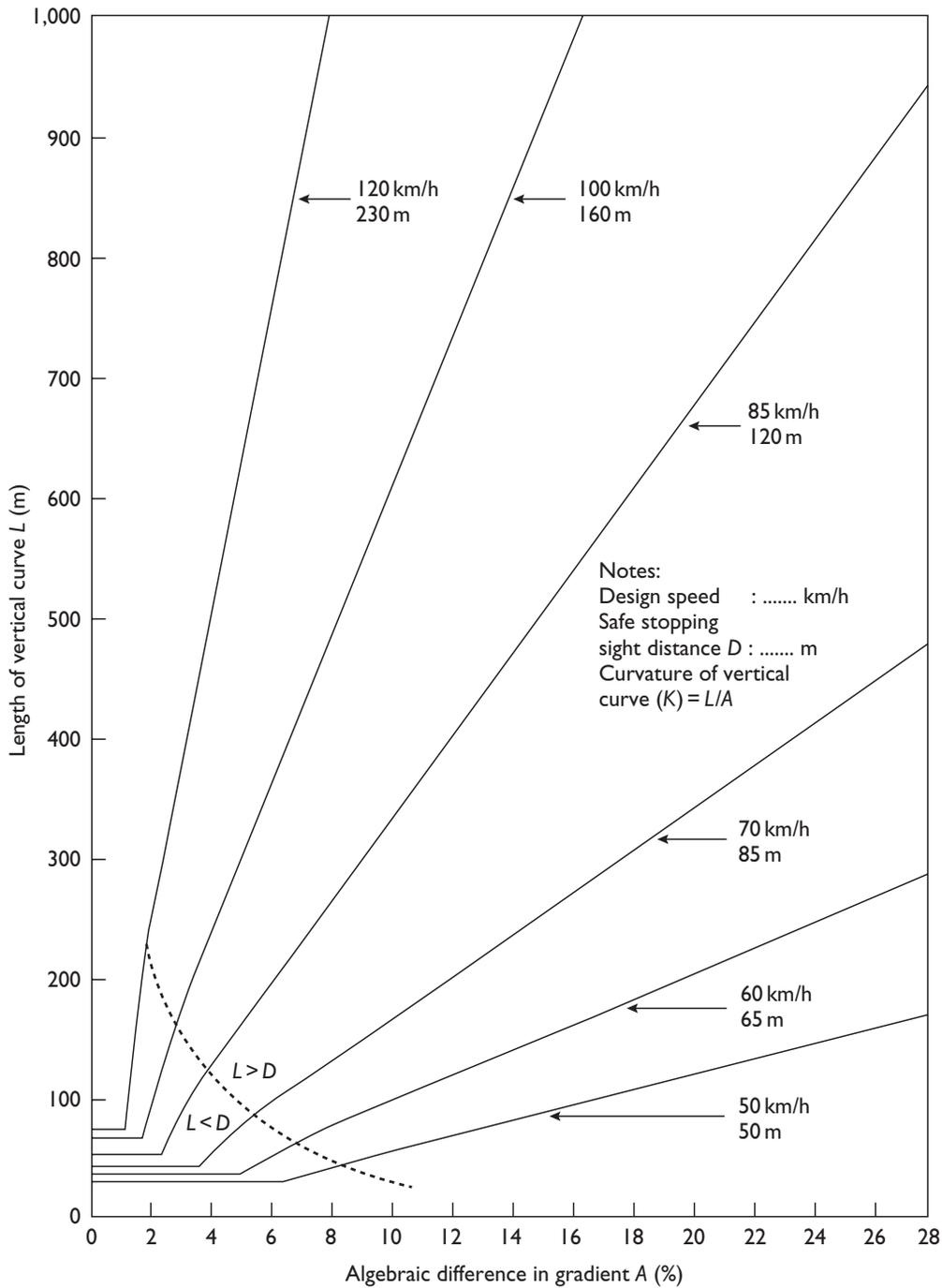


Figure 12.3 Length of crest vertical curves for safe stopping sight distance.
 Assumptions

- Drivers' eye height = 1.05 m
- Object height = 200 mm

Source: TRRL Overseas Unit (1988).

Table 12.5 Minimum levels of acceptable vertical acceleration

Design speed (km/h)	120	100	85	70	60	50	40	30
Vertical acceleration in proportion of 9.81 m/sec ²	0.05	0.06	0.07	0.08	0.08	0.09	0.10	0.10

Source: TRRL Overseas Unit (1988).

For $S > L$:

$$L = 2S - \frac{200(h + S \times \tan \alpha)}{A}$$

where L is the minimum length of vertical sag curve (m); S the required stopping sight distance (m); h the headlight height (m); A the algebraic difference in gradients (per cent); and α the angle of upward divergence of headlight beam (degrees). Appropriate values for h and α are 600 mm and 1.0 degrees respectively.

The use of these equations can lead to requirements for unrealistically long vertical curves as, especially at higher speeds, sight distances may be in excess of the effective range of the headlamp beam, particularly when dipped beams are used. Thus, these equations should only be used on the approaches to fords and drifts, or other locations where water may be present on the road surface.

Comfort
criteria

Thus, sag curves are normally designed on the basis of a driver comfort criterion, using a limitation on the vertical acceleration experienced when traversing vertical curves. The minimum lengths of vertical curve based on this criterion are calculated from the following equation:

$$L = \frac{A \times V^2}{1300C}$$

where L is the minimum length of vertical sag curve (m); A the algebraic difference in gradients (per cent); V the design speed (km/h); and C the vertical acceleration (m/sec²).

Values of C recommended by *Overseas Road Note 6* are shown in Table 12.5. The driver comfort criterion applies to both crest and sag curves. However, on crest curves, this criterion gives much shorter lengths of curve than the stopping sight distance criterion. The minimum lengths of sag vertical curves, based on the driver comfort criterion, are shown in Figure 12.4.

12.4 Phasing

Horizontal and vertical alignment should not be designed independently. Inappropriate combination and mis-phasing of horizontal and vertical curves may present the driver with a confusing view of the road. Such problems often occur on sag curves. Severe phasing problems can be dangerous by concealing hazards on the

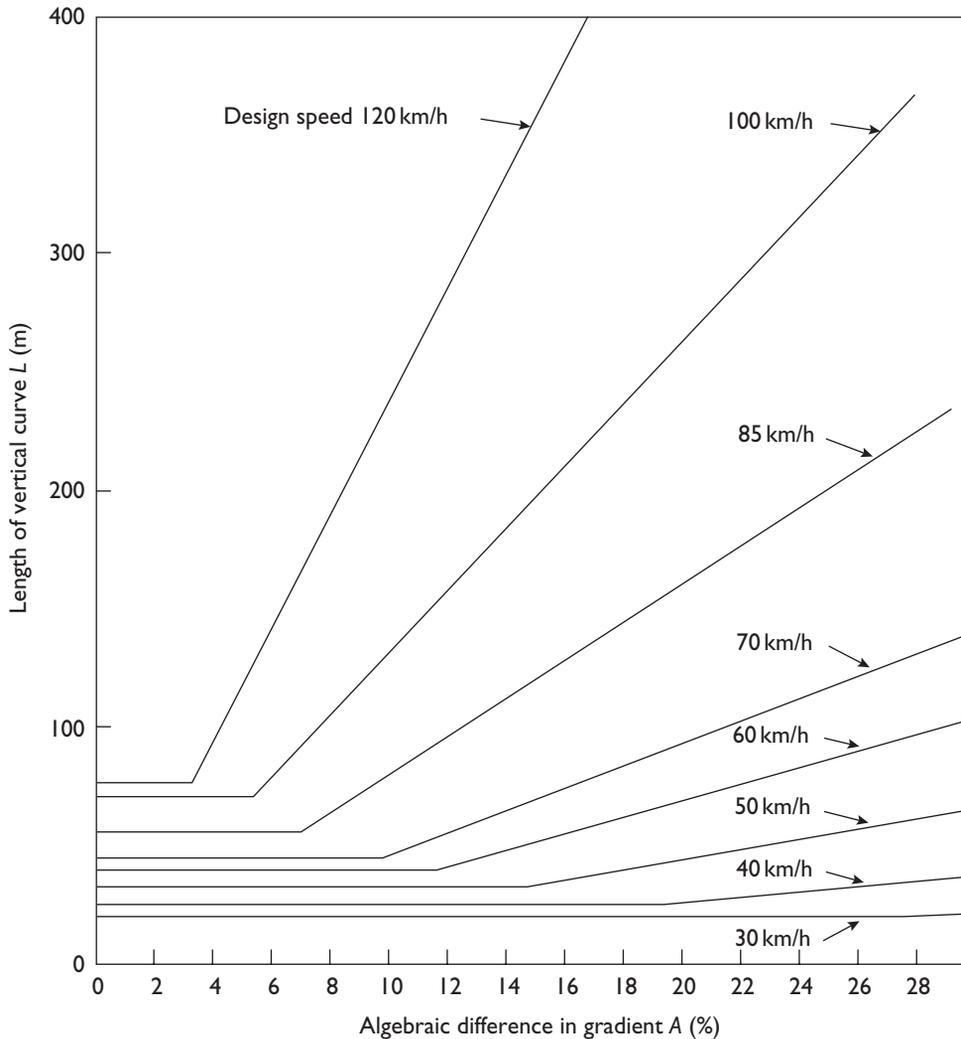


Figure 12.4 Length of sag vertical curve for adequate riding comfort.

Source: TRRL Overseas Unit (1988).

road ahead. The defects become more critical as curve radii reduce. Types of misphasing are described in Box 12.1.

12.5 Alignment selection

12.5.1 Screening

There are several approaches to selecting the route for a road, but all use explicitly or implicitly a 'screening' technique. The first step is to define 'control points'. These are intermediate points of location between the ends of the road that may constrain the route to pass or intersect with a particular feature. They are determined during the reconnaissance study, prior to undertaking any physical surveys.

Control
points

Box 12.1 Types of mis-phasing

Vertical curve overlapping one end of the horizontal curve or located immediately preceding a horizontal curve

A sharp horizontal curve shortly after a pronounced crest curve is a serious hazard that must be avoided. The corrective action is either to separate the curves, apply a more gentle horizontal curve, or make the horizontal curve start well before the summit of the crest curve. In the sag curve situation an apparent kink is produced, creating an aesthetically unpleasant appearance of the road.

Both ends of the vertical curve lie on the horizontal curve

If both ends of a crest curve lie on a sharp horizontal curve, the radius of the horizontal curve may appear to the driver to decrease abruptly over the length of the crest curve. If the vertical curve is a sag curve, the radius of the horizontal curve may appear to increase. The corrective action is to make both ends of the curves coincide, or to separate them significantly.

The vertical curve overlaps both ends of the horizontal curve

If a vertical crest curve overlaps both ends of a sharp horizontal curve, a hazard may be created. This is because a vehicle has to turn sharply while sight distance is reduced on the vertical curve. The corrective action is to make both ends of the curves coincide.

Corridors

Typically, hills, waterways and major transport lines of communications are examined first. For example, it may be necessary to establish an interchange or junction with a major intersecting road or railway, establish a suitable site for a river crossing, or establish low points (or passes) through a range of hills to be traversed.

Having established important control points, the next step is to select broad bands of interest within which to select corridors for more detailed study. A series of traces are identified, within each corridor, indicating where road construction is undesirable for planning and physical reasons. For instance, the planning information trace may identify areas zoned for residential development and/or industrial use; the geological trace may identify areas of peat, marshes, intense gully erosion, areas of potential landslides, mining subsidence etc.; and the drainage trace may identify rivers, streams and canals together with watershed delineation.

Screening factors

Alternative routes should be selected to avoid, as far as possible, the more problematic areas. The following aspects will require particular consideration during the alignment study for rural roads:

- topography;
- in situ soils and availability of road construction materials;
- drainage conditions and need for major structures for waterway crossings;
- neighbouring projects.

The choice between different route options depends upon several considerations, such as the following:

- minimization of construction costs, including earthworks;
- provision of the shortest route between the fixed end points, thereby minimizing user costs;
- achievement of a suitable level of service fulfilling the needs of the road users;
- provision of a safe and environmentally pleasing route alignment, as far as is possible, from all points of view.

A more detailed list of factors is given in Figure 12.5.

Selection of a horizontal alignment along primary or secondary watersheds can usually minimize the construction and maintenance costs. However, this may not be convenient if the road serves agricultural areas, which are often not located in riverine areas. These tend to have weak subsoils, poor construction materials, swamps and flood-prone areas. Often, there is a need for several drainage structures to provide water crossings, including relief culverts in sidelong terrain. Optimal alignment selection requires a value judgement of these conflicting factors to achieve a satisfactory balance.

Watersheds

Topography, physical features and land use have a pronounced effect on road location and geometrics. In rugged terrain, the location and geometric design elements may be governed almost completely by the topography because of limitations imposed by hills, valleys, steep slopes, rivers and lakes. Information regarding topography and physical features should therefore be obtained in the early stages of planning and design.

Topography and physical features

Many costs are proportional to road length so an alignment close to the shortest distance route will be preferred in many situations. The shortest route minimizes pavement construction costs, and operating costs will normally be lower provided that steeper grades are not introduced. If the topography is more gentle along a longer alignment, earthwork quantities can be reduced and lower construction costs can result. The aim will often be to choose an alignment that minimizes total cost.

Shortest distance

During the alignment selection process, the road engineer needs to co-ordinate with a number of different authorities responsible for existing and planned infrastructure. Examples are: other roads, railways, power lines, telecommunication lines, oil pipe lines, water pipe lines, agriculture, irrigation channels, archaeological sites, burial grounds, canals and dam sites, military installations, etc. Co-ordination is often constrained by the lack of information on developments plans in the area. Information on some projects may be classified, for national security reasons, and will be difficult to obtain.

Co-ordination

12.5.2 Sources of information

Topographical maps, with or without contours, are normally available in most countries. The entire country is often covered by maps to scale 1:50,000 while, for certain urban areas, maps at 1:2,500 or 1:5,000 scale may be available. However, maps are often based on 20–40 years old aerial photography and may not be up-to-date with more recent infrastructure developments, such as built-up areas, houses,

Maps

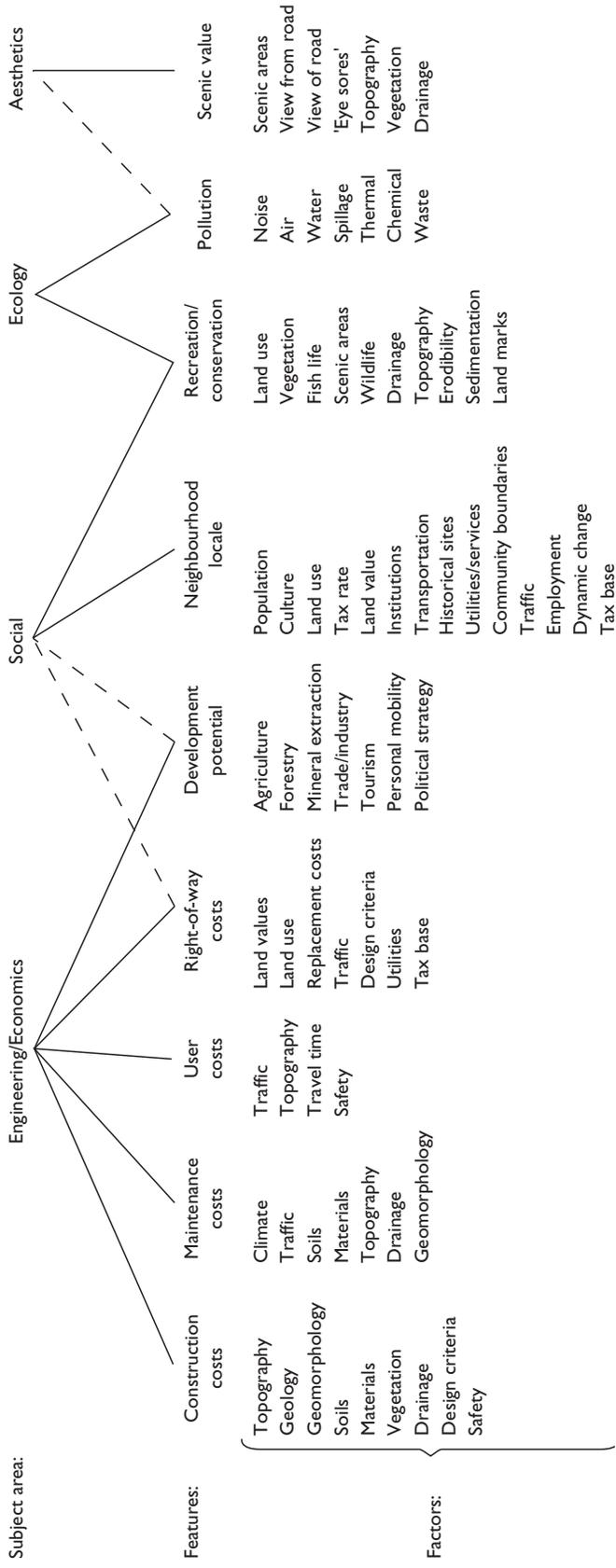


Figure 1.2.5 Factors affecting road alignment selection.
Source: Beaumont and Beaven (1977).

roads, tracks, telephone lines, power transmission lines and land use. Consequently, such maps are normally only suitable for preliminary investigations and reconnaissance purposes, including determination of drainage catchment areas. Geological maps may also be available, and are useful for soils and materials investigations. Should maps not be available in the country, or if they are restricted due to military reasons, they can often be obtained from map publishing bureaux, such as the Institut Geographique National (IGN) in Paris, or international documentation centres in a number of countries. Tactical pilotage charts, or 'pilot maps' are produced for air navigation purposes and are available internationally.

Aerial photographs, which are generally taken at scales ranging from 1:20,000 to 1:60,000, may be very useful to supplement the existing topographical and geological maps. Their prime advantage is in giving a highly detailed, non-interpreted view of the terrain. When studied with the aid of a stereoscope, the ground surface is seen in full three-dimensional relief. Aerial photographs are available in many countries.

Aerial photos

Satellite imagery is available for the whole world apart from persistently cloudy regions. The images depict terrain and drainage systems, and are most useful at the project identification stage of investigation. They also show changes in surface features by repeated coverage of the same area. Change in major river flow patterns, retreating coastlines, deforestation or sand dune movements can be observed in this way. In recent years, satellite imagery, with resolution 1 m or better, has become commercially available (see Chapter 8). These are relatively expensive, and their price is often similar to that of new colour aerial photographs at a scale of 1:25,000, but the price is expected to fall in the future. With 1-m resolution satellite imagery, it is possible to detect the same level of details as in aerial photographs at scales of 1:20,000 to 1:40,000.

Satellite imagery

Topographical maps, aerial photographs and satellite imageries are suitable only for preliminary design of roads. For detailed design, it is necessary to carry out either terrestrial surveys or aerial photography with ground control. Aerial photography is not suitable in areas that have dense vegetation throughout the year. With a flight height of 1,100 m, corresponding to a photograph scale of approximately 1:7,000, maps can be produced with contours having a mean error of 330 mm and a spot level mean error of 220 mm. For roads in rolling or hilly/mountainous terrain, this accuracy is sufficient for road design, including estimation of earthwork quantities for tender purposes. However, for rehabilitation projects with overlays and some improvements in vertical alignment a terrestrial survey is required.

Choice of survey method

12.6 Intersections

Conflicting vehicle movements at junctions are the largest cause of accidents in many developing countries. A small number of well-designed junctions on a route are preferable to a large number of low standard junctions. Simple crossroads have the worst accident record. Staggered crossroads, consisting of two successive T-junctions on opposite sides of the road, can reduce the accident rate. The use of roundabouts, traffic lights and channelization may be appropriate to improve vehicle flow and safety. Conflicts can be largely eliminated by the expensive solution of grade separation, but it is not normally necessary to design for free-flow conditions.

Reducing accidents

Traffic lights	<p>For relatively high traffic volumes, especially in urban areas, it will be more appropriate to improve vehicle flow by traffic light control. Where traffic lights are required, or likely to be installed in the future the intersection spacing should be as regular as possible to permit efficient traffic light phasing on the major road. A guide to the spacing in these conditions is given by:</p> $D = 0.1.39 \times C \times V$ <p>where D is the spacing of intersections (m); C the cycle time (sec); and V the vehicle speed (km/h).</p>
Staggered intersections	<p>Staggered intersections should be used where information on the traffic volumes, or relative importance of two crossing roads, enables the designer to distinguish between the major road and the minor road. Staggered intersections are referred to as either 'left-right' or 'right-left' according to turning movements of a vehicle on the minor road crossing the major road. The staggering should be designed so that traffic does not have to cross the opposing main road traffic stream before exiting. This avoids the need for traffic to wait in the middle of the main road. Consequently, right-left stagger is preferred for left-hand driving, and left-right is preferred for right-hand driving. The minimum spacing between the legs of a staggered intersection should be 100m. Where intersections have tapered deceleration/acceleration lanes, the minimum spacing should be increased to 200m.</p>
Maximum gradient	<p>The longitudinal gradient of any leg of an intersection should not exceed 3 per cent to permit heavy vehicles to accelerate at reasonable rates within the vicinity of the intersection. This maximum gradient should also apply on the intersecting road within 30m of the intersection.</p>
Islands	<p>Kerbed islands can be potential hazards. They should only be considered when their use results in a safety benefit greater than the additional hazard caused by their introduction. In rural areas, channelization with islands in the middle of the road should consist of painted hatched areas (ghost islands) and not by kerbed islands.</p>
Intersection layout	<p>The basic principles of good intersection design are that it should allow transition from one route to another and movement on the main road with minimum delay and maximum safety. Hence, the layout and operation of the intersection should be obvious and unambiguous with good visibility between conflicting movements. These objectives need to be achieved at reasonable cost, so provision of unnecessarily high standards is avoided. Different intersection layouts will be appropriate under varying circumstances depending on traffic flows, vehicle types, travel speeds and site limitations. For feeder roads and access roads with very little truck traffic, the radii for edge of pavement may be 5–7m, as shown in Table 12.6.</p>
Three-arc curve	<p>A three-arc curve conforming to the inner wheel path for the semi-trailer design vehicle is recommended for intersections with arterial roads having considerable truck traffic. The ratio between the radii of the three arcs should be $R1 : R2 : R3 = 2.5 : 1.0 : 5.0$, where $R1$ is the radii of the first curve, $R2$ the middle and $R3$ the radii of the last curve in the direction of turning movement. $R2$ is normally 12m.</p>
Turning lane	<p>Where traffic flows are relatively high, channelization should be established for capacity and safety reasons by provision of a turning lane on the main road and an island on the side road, as in Figure 12.6.</p>

Table 12.6 Intersection layout standards

Intersection type	Applicable for intersection with	Truck traffic (per day) on side road	Carriageway width of side road (m)	Radii of pavement edge curve (m)
A	Principal arterial	> 100	6.5–7.5	3-arc curve
B	Minor arterial	40–100	6.5–7.0	15 m or 3-arc curve
C	Collector road	10–40	5.5–6.5	10 m
D	Feeder road	< 10	3.0–5.5	7 m
E	Access road	< 1	3.0	5 m

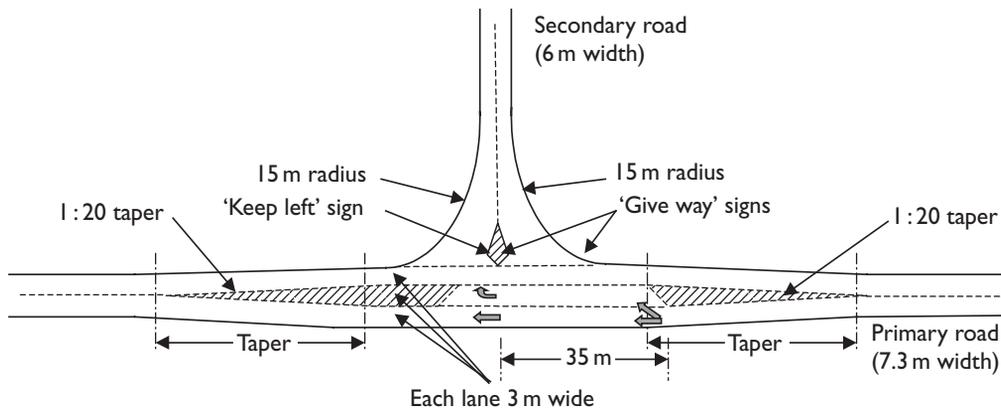


Figure 12.6 Dimensions for typical priority T-junction (driving on the left).

Adapted from: Department of Transport (1987).

12.7 Low-cost roads

Low-cost roads are characterized by:

- design year traffic of less than 20 vehicles per day;
- gradients constructed generally close to the ruling grade of the natural ground;
- constructed mainly from the in situ soils, with no significant quantity of imported manufactured materials;
- constructed on the basis of minimal designs.

Designing and constructing low-cost roads using techniques similar to those for major highways can increase engineering costs to an unacceptable proportion of total costs. Field surveys and the extent of design works should therefore be tailored to reflect the low-cost nature of the works.

Low-cost roads are normally constructed by rehabilitation/upgrading of existing tracks. Their design can therefore be carried out using a simple road location plan compiled from existing 1:50,000 topographical maps. The plan can show the

Characteristics

Location plan and works schedule

approximate position of the horizontal alignment, indicating stations and the location of existing and potential gravel borrow pits and stone quarries. The design information can be specified on a simple works schedule, which outlines for each cross-section:

- cross-section type;
- formation – whether light reshaping (say less than $2\text{ m}^3/\text{m}$) or heavy reshaping ($2\text{--}5\text{ m}^3/\text{m}$) is needed;
- compacted thickness of fill (m);
- excavation (m^3) – either common excavation, or rock excavation;
- compacted thickness of gravel wearing course (mm);
- drainage requirements for each side of the road – mitre drain spacing; catchwater drains;
- erosion checks for each side of the road;
- drainage structures – whether repair of existing structures is necessary, or whether a new structure is needed;
- other notes and remarks.

Horizontal alignment

Horizontal alignment data are not normally needed, but possible alignment improvements may need to be indicated in the works schedule, and decided finally during setting out on site.

Vertical alignment

The vertical alignment of low-cost roads will generally follow the existing ground. Therefore, it is not normally necessary to prepare a design for the longitudinal profile. The earthworks required to form up the roadbed are normally quite modest, except on hilly sections of road where cut and fill operations require longitudinal haulage of material. More significant cut or fill sections may be needed where improvement of sight distance over crests is required, where the road needs raising over water crossings, or where the road needs to be raised due to the natural soil having low bearing capacity or in areas liable to flooding. Earthworks will only need to be measured accurately in these areas of relatively significant cut and fill operations. In other areas, where standard cross-sections are used, earthworks can be calculated on the basis of a metre-of-road global cost. This reduces to a minimum the need for levelling and other time- and people-demanding surveys, thereby reducing costs and speeding-up construction.

12.8 Computer-aided design

Increasingly, road administrations throughout the world make use of computer-aided design (CAD) for their road works, including all project documents and drawings. A number of CAD platforms are used, including 'Moss', 'AutoCad' and 'MicroStation'. Other systems include fully integrated computer-aided engineering (CAE) platforms. These include 'InRoads' and 'NovaCad'. CAE road design programs offer

- Direct transfer and processing of topographical survey data, including the development of digital terrain models (DTM) and the production of contoured maps.

- Design of horizontal alignment, possibly with transition curves, and vertical alignment design including production of plan and profile drawings.
- User-defined typical cross-sections, including superelevation, and slope and ditch control by parametrically driven decision tables.
- Earthwork calculations, which enable optimization of earthwork quantities using mass-haul curves.
- Three-dimensional visualization.

References

- Department of Transport (1987) *Roads and Traffic in Urban Areas*, London: H.M. Stationery Office.
- Beaumont, T.E. and Beaven, P.J. (1977) 'The use of satellite imagery for highway engineering in overseas countries', *TRRL Supplementary Report 279*, Crowthorne: Transport and Road Research Laboratory.
- Botswana MoWC (1982) *Botswana Road Design Manual*, Gaborone: Ministry of Works and Communications, Botswana.
- SATCC Technical Unit (1995) *Recommendations On Road Design Standards, Volume I: Geometric Design of Rural Roads*, Maputo: Southern Africa Transport and Communications Commission.
- Snowy Mountains Engineering Corporation (undated) *Road Design Manual*, Thimphu: Public Works Department, Royal Government of Bhutan.
- TRRL Overseas Unit (1988) 'A guide to geometric design', *Overseas Road Note 6*, Crowthorne: Transport and Road Research Laboratory.

Earthworks, unbound and stabilized pavements

Bent Thagesen

13.1 Earthworks

13.1.1 Activities

Construction of new roads, especially major highways, nearly always involves the movement of soil and rock prior to building the road pavement. General aspects of earthwork operations are described in many textbooks. Some special features of earthworks in tropical rocks and soils are described in *Overseas Road Note 31* (TRL 1993) and by Millard (1993). Earthwork operations may be classified as:

- Clearing and grubbing
- Excavation
- Construction of embankments
- Compaction
- Finishing operations.

13.1.2 Clearing and grubbing

Clearing and grubbing involves the removal of trees, stumps, roots, debris, etc. from the area of proposed excavation and embankment. Clearing refers to the removal of material above the existing ground surface. Grubbing is the removal of objects to a nominal depth below the surface. On equipment-based road projects, clearing and grubbing operations are generally performed by bulldozers with various attachments (see Figure 13.1). A considerable amount of hand labour may also be necessary.

13.1.3 Excavation

Excavation is the process of loosening and removing rock or earth from its original position and transporting it to a fill or waste deposit. Excavation is often divided into three categories: road and drainage excavation, excavation for structures and borrow excavation.

Road and
drainage

Road and drainage excavation involves the removal of material from cuttings and the excavation of ditches. Excavated materials capable of being compacted to form a stable fill are used for construction of embankments, subgrades and shoulders, or

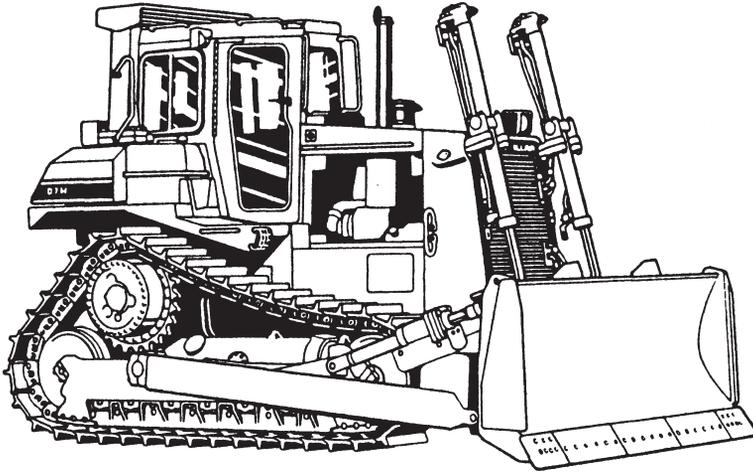


Figure 13.1 Bulldozer.

as backfill for structures. Unsuitable and surplus excavated material is disposed of as 'spoil'. Excavated topsoil is usually stockpiled for later use on side slopes in cuttings and on embankments. If the soil forming the bottom of a cutting is not suitable as the foundation for the road pavement, it may be necessary to remove the soil and replace it with satisfactory material.

It is important that all cuts are kept well drained during the whole excavation operation to prevent wetting and softening of the soil. Ditch work should be scheduled with this object in view.

Cuttings through sound rock can stand close to vertical. In weathered rock and soil, conditions are more unstable. Usually, side slopes are made steeper than the original ground surface, and slides may occur during construction or at a later date. In some mountainous countries, even natural ground slopes may be unstable, and landslides are a persistent problem. An example is the Himalayan Kingdom of Bhutan where, in 1993, more than 130 slides were 'live' along the 400-km road from the capital Thimphu to Mongar in the east of the country.

Instability of natural and construction slopes may be indicated by a range of features:

- springs and patches of reeds in the slope side (indicating that water may cause problems during the rainy season);
- trees leaning at different angles (not trees leaning downslope at the same angle);
- an irregular, pocketed surface with cracks and small ponds, or the presence on the slope of hollow bowl-shaped depressions with a steep head;
- weak, weathered and highly fractured rocks;
- rocks with bedding lying parallel to the hillside.

Slides are usually provoked by an accumulation of water in the rock or soil, and often occur along planes of weakness. The stability of slopes depends on many factors,

Drainage

Slope stability

Features of instability

the most important being the dimension of the slope, the angle of the slope, the type of material, and stratification of the rock or soil. Investigations of difficult situations should be left to specialists.

Reducing the risk of slides

The risk of slides in cuttings can be reduced by flattening all slopes but, in steep and mountainous terrain, this may not be practicable. Even in cases where the slopes can be flattened, the most economical solution may be to adjust the slopes only when and where slides occur. If possible, new roads should not be located in areas where slides can be expected. Construction of roads in mountainous regions is discussed in *Overseas Road Note 16* (TRL 1997).

Equipment

Selection of equipment for excavation and moving of earth depends on the nature of the material, how far it has to be hauled, climatic conditions, the skill and knowledge of the equipment operators and economic considerations. Excavation and moving of earth can be carried out through a wide range of methods using various mixes of labour and equipment (see also Chapter 18). On equipment-based projects, bulldozers are normally used for very short hauls and for spreading dumped materials. For moderate and longer hauls scrapers push-loaded by tractors may offer lower costs. However, this type of very specialized and expensive equipment is rarely seen outside the United States. For hauls of considerable length, or over the public highways, trucks loaded by front-end loaders (see Figure 13.2), or power shovels may be the cheapest solution. A motor grader (see Figure 13.3) may be used for shaping ditches, trimming slopes and shaping the cross section of the road. A crane with



Figure 13.2 Front-end loader.

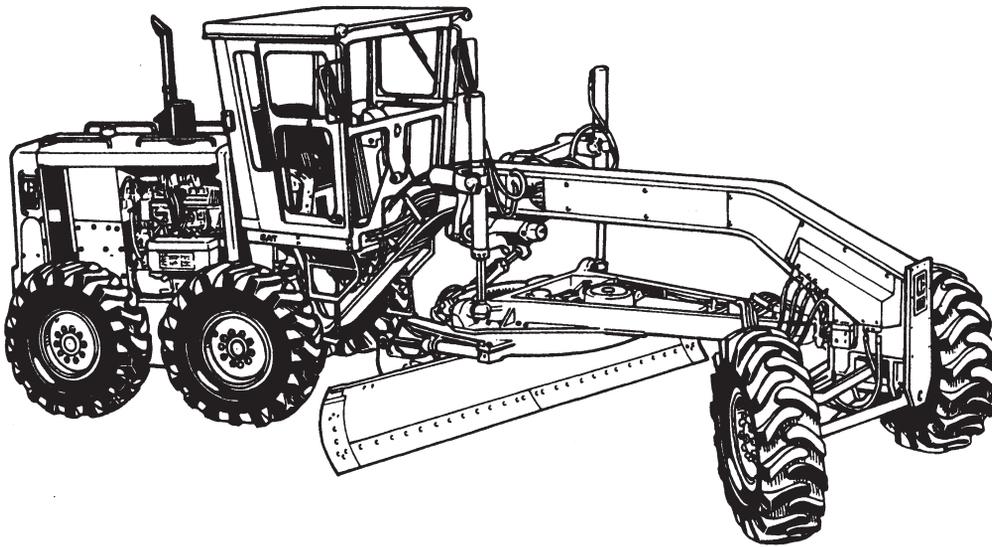


Figure 13.3 Motor grader.

a dragline is frequently used for excavation in swampy land. In the case of rock excavation, drilling and blasting will be necessary, and a crane with shovel may be needed to load the blasted rock into trucks.

Structural excavation includes the excavation of material to permit the construction of culverts, foundations for bridges, retaining walls, etc. Suitable materials taken from structural excavations are used either in backfilling around the completed structure or in other parts of the construction site. Both machine and hand methods are used for structural excavation, but more hand labour is required in this operation than in other types of excavation. Bulldozers can be used to good advantage for the excavation and backfilling of many structures. Front-end loaders also find application in this work. For deep excavation, a crane with a clamshell is suitable because of its ability to excavate vertically and its close proximity to the bracing and sheeting required in deep excavations.

Structural
excavation

Sometimes sufficient material for the formation of embankments is not available from excavations performed within the limits of the right-of-way. In such cases, additional suitable material is taken from borrow pits. Specifications for borrow excavation are generally the same as for the main works. Additional requirements usually relate to environmental considerations concerning the condition in which borrow pits should be left when they are abandoned. Borrow-pit excavation may be carried out using shovels or front-end loaders, and the material loaded and transported on trucks.

Borrow
excavation

13.1.4 Construction of embankments

Embankments are used in road construction when the vertical alignment of the road has to be raised above the level of the existing ground to satisfy design standards,

or prevent damage from surface or ground water. Many embankments are only 0.5–1.5 m high, but heights of 5 m or more may be used on major highways.

Settlement

High embankments impose a heavy load on the underlying foundation soil. On some soils, this may result in settlements. If the foundation soil is extremely weak, a slip failure may occur. The residual soils, common in many tropical countries (e.g. laterite), are not normally very compressible. Any small settlement that does occur will have stopped before the embankment is completed. An exception to this is with clays developed by weathering of volcanic ash (see Chapter 9). Some transported soils are also very prone to settlement. Examples are windblown sands and alluvial clays including organic marsh soils.

Improving stability

The stability of the foundation for embankments depends on the dimension of the embankment, and the type and thickness of the foundation soil. Expert advice should be obtained when dealing with difficult situations. Methods for improving the structural properties of an embankment foundation include the removal and replacement of unsuitable material. Consolidation (extrusion of water) of saturated, fine-grained soils can be accelerated by use of surcharge and/or vertical sand drains. However, these procedures are very costly, and relocation of the route should always be considered whenever extremely weak soils are encountered at considerable depth.

Side-sloping terrain

The cross-section of a road embankment in flat terrain usually consists of a flat top section with symmetrical side slopes on either side. In side-sloping terrain, it is common practice to cut the road partly into the hillside and use the excavated material as fill on the outside of the road. The result is an asymmetric cross-section as illustrated in Figure 13.4. In steep, side-sloping terrain, the embankment is usually keyed to the hillside by cutting horizontal benches into the slope beneath the fill. A retaining wall may be needed to support the fill. The material has to be adequately compacted and well drained.

Suitable fill

Most types of soil and broken rock can be used for construction of embankments, but materials of the American Association of State Highway Officials (AASHTO)

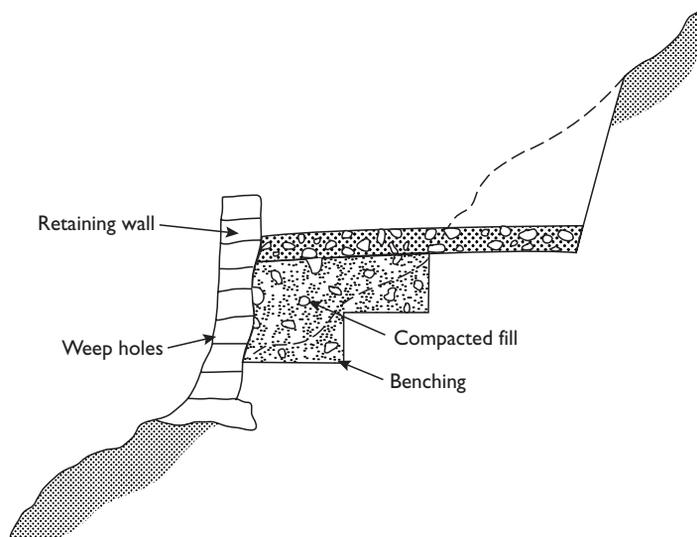


Figure 13.4 Cross-section of a road in steep side-sloping terrain.

classification A-1, A-2-4, A-2-5 and A-3 are preferred. More plastic materials may create problems in wet weather. If expansive clays cannot be avoided, special precautions should be taken, as described in Chapter 9. Highly expansive clays and organic soils should not be used as fill.

Embankments made from materials of low plasticity can be constructed with slopes as steep as 1 : 1.5 (vertical : horizontal) without causing slope stability problems. For other soil types, particularly in wet climates, the maximum slope should be 1 : 2. Where the embankment is subjected to flooding, the slope should not be steeper than 1 : 3. Side slopes are normally protected from the erosive action of wind and water by establishing a cover of vegetation.

Earth embankments are constructed using relatively thin layers of soil. On equipment-based projects, the material is usually dumped at the required location by trucks, and spread by bulldozers or graders. The maximum thickness of loose soil is usually 250–300 mm. The soil is thoroughly compacted before the next layer is placed. During the construction operation, the embankment should be kept well drained at all times, especially when material of high plasticity is used.

13.1.5 Compaction

Compaction increases the density of a material by expelling air from the voids in the material and, thereby, bringing the particles into more intimate contact with each other. Compaction is the cheapest and simplest method for improving the shearing resistance of soil and minimizing future settlements. Therefore, soils in embankments and subgrades in cuttings are usually compacted using special compacting equipment, such as rollers, vibrators or tampers. The result of compaction work depends primarily on the moisture content of the soil, the type of the soil, the compaction equipment used and the energy applied.

For most soil types, the maximum dry density is achieved at a particular moisture content, the so-called optimum moisture content (see Chapter 8). Because of the moisture–density relation, water must be added to dry soils, and overly wet soils must be aerated before compaction. However, in dry areas, where it is difficult to provide the large amount of water needed to bring the moisture content of a dry soil to the optimum level, it may be better to compact the soil in the dry state rather than adding an insufficient amount of water to the soil. In rainy weather, it may be necessary to replace an overly wet soil with more suitable material, or to stabilize the wet soil with lime (discussed later).

The greater the permeability, the easier it is to expel air from the soil. This partly explains why gravel and sand can usually be compacted to a greater density than clay. In well-graded soils, the smaller particles may fill some of the voids between the larger particles. Thus, well-graded soils can achieve a greater density than uniformly graded soils.

The choice of compaction equipment is wide. Most equipment is available in several sizes (deadweights). Compactors may be self-propelled or pulled by a tractor. The most common type of compactor is the self-propelled, smooth, steel-wheeled roller. The slow-moving roller acts on the soil with ‘static’ loads. It can be used for compacting all types of soil, that is, gravel and sand as well as clay. The steel-wheeled roller may be equipped with devices for vibrating the wheels.

Slopes

Construction

Purpose

Moisture

Soil type

Equipment

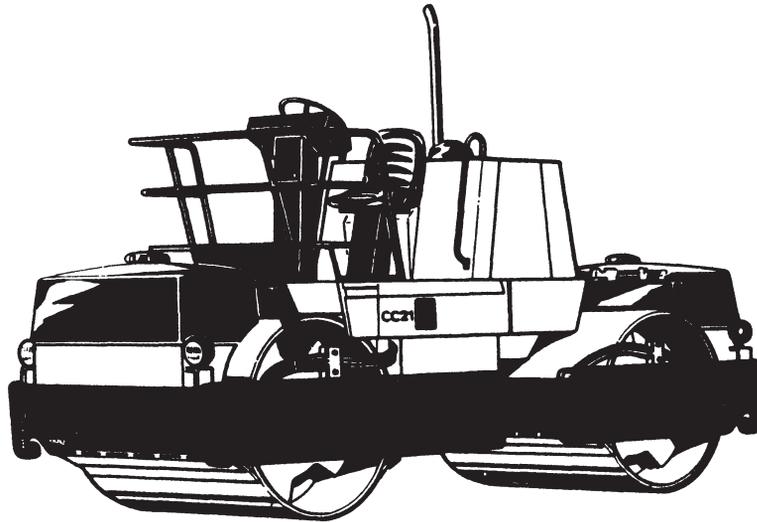


Figure 13.5 Self-propelled smooth steel-wheeled roller.

The vibration imposes pulsating stresses in the soil. This reduces the friction between the soil particles and results in a highly effective and deep compaction, particularly in gravel and sand. If the vibration is turned off, the roller can be used as a traditional, static roller. This makes the vibratory steel-wheeled roller (see Figure 13.5) very versatile.

Other types

Other types of compactor include the pneumatic-tyred roller and the sheepsfoot roller. A pneumatic-tyred roller consists of rubber-tyred wheels mounted on an articulated frame, which provides a uniform load on each wheel. A sheepsfoot roller is made of a steel drum to which protruding, tamping feet have been attached. For compaction of backfill in narrow trenches and excavations for structures, a wide variety of hand-operated mechanical tampers and plate type vibrators are available.

Energy

It is possible to compact a soil to a higher density and to a greater depth by increasing the loading per unit width of the roller (see Figure 13.6). However, the loading must not exceed a critical limit that depends on the soil type. If the roller produces shear stresses exceeding the shear strength of the soil, the soil will be remoulded and loosened. As seen from Figure 13.6, repetitions of roller passes are effective only up to a certain limit, in most cases. The compaction effect usually fades out when the number of passes exceeds 8–16, depending on the soil type.

Compaction requirements

Compaction requirements are commonly specified using an end-product specification. Compaction tests on the particular soil are carried out in the laboratory, and the results are used to define the required density. During construction, the densities obtained in the field are determined and compared to that required. The most common methods applied for measuring field densities are the sand cone and the nuclear density gauge. It is normally recommended that the top of an embankment, as well as the upper 500mm of the subgrade in cuttings are compacted to minimum 95 per cent of the maximum dry density obtained in the modified proctor test. For the

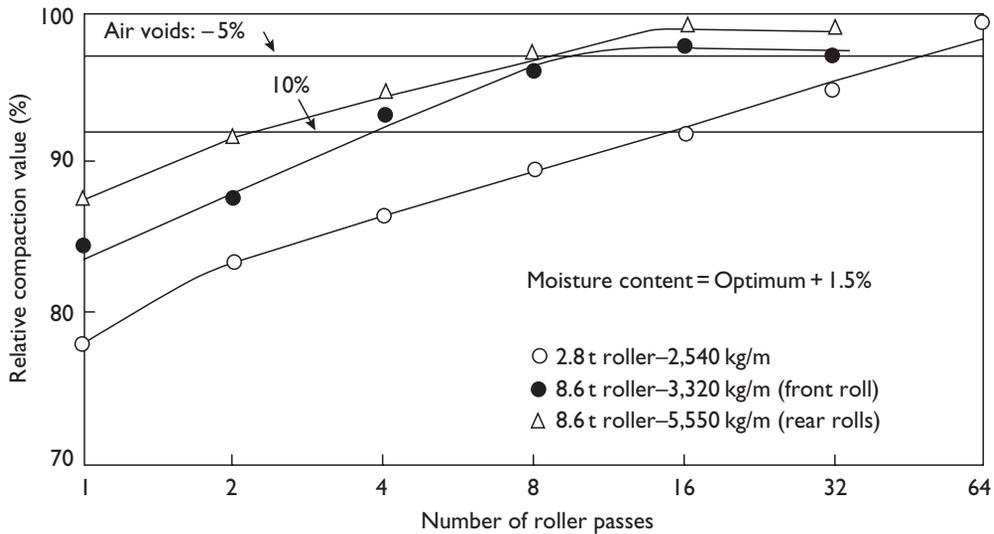


Figure 13.6 Relationship between relative compaction of a sandy clay and the number of passes of smooth steel-wheeled rollers.

Source: Parsons (1992).

lower layers of an embankment, the requirement is reduced to 93 per cent. For coarsely grained soil material, the proctor test should be replaced by vibro-compaction.

Most soil materials change volume when compacted. Rock fill will occupy 1.2–1.5 times more space than the solid rock. Excavated earth will expand in the transporting vehicle but, when compacted, it will usually shrink to 0.6–0.8 times the original volume, depending on the soil type. It is important to take this volume change into consideration when computing excavation, haulage and fill quantities.

Volume
change

13.1.6 Finishing operations

Finishing operations are the final activities necessary to complete the earthwork. These involve trimming of formation level, shoulders, ditches and side slopes. Most finishing operations are carried out concurrently with other earthwork operations and performed as the job approaches completion. The equipment most widely used for finishing is the motor grader and the bulldozer.

13.2 Pavement structure

A road pavement is a structure whose primary aim is to support the traffic loads and transmit them to the basement soil – after reducing the stresses to below the level that can be supported by the soil. Most pavements consist of three superimposed layers, each performing different primary functions. The traffic-induced stresses and the influence from weather decrease with increasing depth in the pavement, as do the quality requirements and the costs of the pavement materials. The terms used in this book for the different layers in a pavement are shown in Figure 13.7.

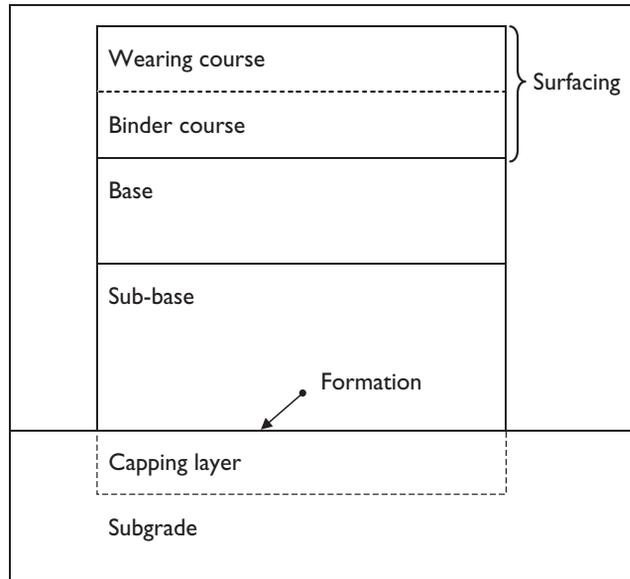


Figure 13.7 Terms used for the different pavement layers.

Wearing course

The wearing course is the uppermost layer of a sealed pavement. It provides the riding surface for the road users. The wearing course should be smooth and dust-free and have adequate skid resistance. It should protect the pavement against wear from the traffic and soaking from rainwater. The most common materials for wearing courses are bituminous surface dressing and asphalt concrete. Earth roads and gravel roads have no wearing course, by definition.

Binder course

A premixed asphalt surfacing is sometimes laid in two layers of different materials. The lower layer of such a construction is known as the binder course. The wearing course and binder course make up the 'surfacing' of the pavement.

Base

The base is the main load-spreading layer. Construction of the base extends some distance beyond the edge of the wearing course to make sure that underlying layers will support the loads applied at the edge of the pavement. The base may consist of premixed asphalt, cement concrete, graded granular gravel, crushed rock, macadam or materials stabilized with lime or cement. Cement concrete roads combine the base and surfacing in one layer. Cement concrete pavements are beyond the scope of this book.

Sub-base

The sub-base is a secondary load-spreading layer. It acts as a separation between the subgrade and the base, and provides a working platform during the construction of the upper layers in the pavement. When the sub-base is made from unbound materials, it may also function as a filter and drainage layer. The sub-base is usually constructed from natural gravel or from materials stabilized with cement or lime. Many pavements have no sub-base layer, and the base is placed directly on the foundation.

Subgrade

The subgrade is the soil acting as a foundation for the pavement. The subgrade is the result of the earthwork, and it may consist of the undisturbed, local soil or material excavated elsewhere and placed as fill. The surface of the subgrade is called the 'formation'.

A capping layer is a strengthening layer that may be used on very weak subgrades. It is normally defined as part of the subgrade. It is made from imported, selected fill or from subgrade material stabilized with lime.

Capping layer

It should be noted that different terminology is used for the pavement layers in the United Kingdom. Binder course and base are called 'base course' and 'road base', respectively.

Other terminology

13.3 Unbound pavement layers

13.3.1 Components

Unbound pavement layers include the capping layer, sub-base and base, constructed from gravel and stone materials without the addition of a binder. *Overseas Road Note 31* (TRL 1993) gives guidance on the construction of these layers and the criteria that should be met by the materials.

Layers

If possible, unbound pavement materials (except dry-vibrated macadam) should be wet during transport and laying to reduce the likelihood of particle segregation. The materials are normally spread by hand or with a bulldozer and/or a motor grader, but use of a spreader is preferable. After spreading, the layer is compacted to specified density using static or vibrating steel-wheeled rollers. The moisture content must be kept as close as possible to the optimum for compaction with the particular roller in use to achieve the specified density. In dry areas, where water is scarce, dry compaction may be used.

Construction

13.3.2 Capping layer

Unbound capping layers are normally made from gravelly soils. A minimum california bearing ratio (CBR) of 15 per cent is recommended for material compacted to the specified density, usually 95 per cent of the maximum dry density obtained with modified (heavy) proctor compaction. The samples should be tested at the highest moisture content expected to occur in the field.

13.3.3 Sub-base

Unbound sub-bases are generally made from naturally occurring gravel. Screening may be necessary to remove larger size particles. Recommendations for the particle size distribution are given in Table 13.1.

Grading requirements

The fines of granular sub-base materials should have a limited plasticity, depending on the moisture regime where the material is used. The liquid limit (LL) and the plasticity index (PI) should not exceed the values shown in Table 13.2. If the material is too plastic, it may be modified by mixing it with a little cement or hydrated lime.

Plasticity

A minimum CBR of 30 per cent is recommended for sub-base materials compacted to required field density, normally 95 per cent of the maximum dry density using modified proctor compaction. Given good drainage and a deep-lying water table, the samples should be tested at the optimum moisture content. If saturation of the sub-base is likely to occur, the samples should be tested after four days of

CBR

Table 13.1 Recommended particle size distribution for granular sub-base materials

Sieve size (mm)	Passing sieve (% by mass)
50	100
37.5	80–100
20	60–100
5	30–100
1.18	17–75
0.3	9–50
0.075	5–25

Source: *Overseas Road Note 31* (TRL 1993).

Table 13.2 Maximum plasticity values for sub-base materials

Climate	Liquid limit (%)	Plasticity index (%)
Wet tropical	35	6
Seasonally wet	45	12
Arid and semi-arid	55	20

soaking. For coarse graded material, it may be more appropriate to replace the proctor test with vibro-compaction and omit the CBR test.

Filter criteria

If the sub-base has to protect a drainage layer from blockage by finer materials or prevent mixing of two layers, then additional filter criteria must be met.

13.3.4 Base

13.3.4.1 Materials used

Unbound bases may be constructed from a range of different granular materials. The most widespread types are

- Natural gravel
- Crushed gravel and crushed rock
- Dry-vibrated macadam.

13.3.4.2 Natural gravel base

Grading requirements

Table 13.3 shows the recommended particle size distributions for natural gravel suitable for base construction. The 10-mm nominal size should only be used on roads that carry light traffic. To meet the grading requirement, screening and crushing of larger particle sizes may be required. All grading analyses should be made on materials that have been compacted to simulate breakdown of weak particles under construction. For materials whose stability decreases with breakdown, it may be necessary to specify a maximum los angeles abrasion value or maximum aggregate impact value.

The fines of the gravel should preferably be non-plastic. The PI should not exceed 6 per cent. In arid and semi-arid areas, the maximum allowable PI can be increased to 12 per cent. In cases where it is difficult to meet the plasticity requirement, the gravel may be modified by mixing with a few per cent of cement or hydrated lime.

A minimum CBR of 80 per cent is recommended for base materials after being compacted to specified field density. This is normally 98 per cent of the maximum dry density achieved in the modified proctor compaction test. All samples should be tested after four days of soaking. In arid and semi-arid areas, the specified minimum soaked CBR is reduced to 60 per cent. Coarse gradings may be unsuitable for proctor and CBR tests.

In many countries, natural gravels that do not meet the normal requirements, are used successfully for base construction. They include lateritic gravels with gap-graded particle size distribution, calcareous gravels with high plasticity, and volcanic gravels with low particle strength. The use of these materials should be encouraged if local experience has proved it to be viable. However, the use of 'marginal' materials should normally be confined to roads carrying light traffic.

Granular materials derived from weathering of basic igneous rocks, such as basalt and dolerite, may release detrimental plastic fines during construction and future service (see Chapter 9). No single test method is able consistently to identify unsuitable materials in this group, and expert advice should be sought when considering their use.

Plasticity

CBR

Marginal materials

Weathered rock

13.3.4.3 Crushed gravel and crushed rock base

Crushed gravel is produced by crushing and screening natural gravel and boulders. A proportion of the fine fraction may be natural, uncrushed material. Crushed rock is manufactured by crushing fresh, quarried rock. Recommended grading limits are shown in Table 13.4. The particles should be angular in shape, with a flakiness index less than 35 per cent. Specific limits on the maximum los angeles abrasion value or the aggregate impact value may be used to ensure adequate durability.

Grading requirement

Table 13.3 Recommended particle size distribution for natural gravel base materials

Sieve size (mm)	Passing sieve (% by mass)		
	Nominal maximum particle size (mm)		
	37.5	20	10
50.0	100	—	—
37.5	80–100	100	—
20.0	60–80	80–100	100
10.0	45–65	55–80	80–100
5.0	30–50	40–60	50–70
2.36	20–40	30–50	35–50
0.425	10–25	12–27	12–30
0.075	5–15	5–15	5–15

Source: Overseas Road Note 31 (TRL 1993).

Table 13.4 Recommended particle size distribution for crushed gravel and crushed rock base materials

Sieve size (mm)	Passing sieve (% by mass)		
	Nominal maximum particle size (mm)		
	37.5	28	20
50.0	100	—	—
37.5	95–100	100	—
28.0	—	—	100
20.0	60–80	70–85	90–100
10.0	40–60	50–65	60–75
5.0	25–40	35–55	40–60
2.36	15–30	25–40	30–45
0.425	7–19	12–24	13–27
0.075	5–12	5–12	5–12

Source: *Overseas Road Note 31* (TRL 1993).

Plasticity The PI should not exceed 6 per cent. It may be necessary to add a little cement or hydrated lime to meet this requirement. Great care should be taken when dealing with materials originating from weathering of basic igneous rocks.

CBR The compaction requirement is usually the same as for a base made from natural gravel; that is, 98 per cent of the maximum dry density obtained in the modified proctor test. When properly constructed, crushed gravel and crushed rock base will have CBR values in excess of 100 per cent and there is no need to carry out CBR tests.

13.3.4.4 *Dry-vibrated macadam bases*

Grading requirements Dry-vibrated macadam consists of single-sized crushed stones and fine gravel. The recommended nominal size of the stones is 37.5 or 50.0 mm. The fine gravel should be well-graded material passing the 5.0 mm sieve. In the United Kingdom, this material is called ‘dry-bound macadam’, which is a little misleading as no binder is added to the material.

Construction Construction of dry-vibrated macadam involves spreading of stones in a layer of compacted thickness not more than twice the stone size. The layer is compacted with a steel-wheeled roller. Then, dry fine gravel is spread onto the surface and vibrated down into the cavities between the stones. Any surplus material is brushed off and final compaction carried out. This sequence is repeated until the design thickness is achieved. The method is particularly suitable in dry regions, if stone material is readily available and water is scarce.

13.4 Design of gravel pavements

In emerging and developing countries, a substantial part of the road network is surfaced with earth or gravel. Most rural access roads are likely to remain unsealed for many years, and many new gravel roads will be needed in the future.

Normally, the thickness of a gravel surfacing is not designed, but based on experience. Gravel is typically placed in a compacted thickness of 150–200 mm. In areas where the thickness appears inadequate, resulting in rapid formation of deep ruts, more gravel is added.

Gravel
thickness

New gravel roads are often constructed along the alignment of old tracks or earth roads. Before upgrading an old track or earth road, careful attention should be given to the drainage. Alignments following valley bottoms, subject to flooding during the wet season, should be relocated to higher levels. Roadside ditches, turnouts and culverts should be provided where necessary. There are many examples where new gravel roads have been washed away during the first rainy season due to missing or inadequate drainage.

Drainage

Gravel roads are dusty in the dry season and need frequent maintenance to shape them, repair pot-holes and replace the gravel. When traffic reaches a certain level, the question of upgrading and providing a bituminous seal arises. An investment model can be used to evaluate the stage at which it is economical, in whole life cost terms, to upgrade a road (see Chapter 21). In areas where there is a shortage of gravel, sealing can be economical at quite low traffic levels. Indeed, there are many who argue that if social costs and benefits are treated properly in the appraisal, then upgrading can be economical under a wide range of circumstances.

Sealing of
gravel roads

13.5 Stabilized pavement layers

13.5.1 Stabilization

Stabilization is a process by which a soil is improved and made more stable. The most widely used method of stabilization is ‘mechanical stabilization’, where the material is made more stable by adjusting the particle size distribution. The properties of a soil material may also be improved by addition of a binder. Sandy soils are sometimes stabilized by mixing the soil with emulsified bitumen. Enzyme-based stabilizing agents are increasingly used for improving the properties of organic soils, but the most widely used binders for soil stabilization are portland cement and hydrated lime. Sherwood (1993) gives a detailed account of soil stabilization with cement and lime. Recommendations for the criteria that should be met by materials used in the tropics are published in *Overseas Road Note 31* (TRL 1993).

13.5.2 Cement stabilization

Portland cement is produced by heating to high temperature a mixture of calcareous materials (such as limestone), and materials containing silicates and aluminates (such as clay or shale). In the presence of water, portland cement forms hydrated calcium silicates and aluminates which, in time, form a hard matrix in which mixed aggregate are embedded. Cement is produced all over the world, including most emerging and developing countries.

Cement

In theory most soil materials, with the exception of organic soils and soils containing sulphates, can be stabilized with cement. However, it is difficult to mix cement intimately with clayey materials, and soils with high plasticity are usually best treated with lime. Uniformly graded materials are normally too costly to stabilize with

Soil materials

cement, as they require a large cement content and are difficult to compact. Also, very silty materials may prove difficult because they are sensitive to small changes in moisture content. It is desirable that materials to be stabilized with cement have a PI of less than 20 per cent and a minimum coefficient of uniformity of five.

Cement content

The quality of cement-stabilized materials is usually assessed on the basis of the unconfined compressive strength test on compacted samples that have been allowed to harden for a specified time period. Normally, the strength increases linearly with the cement content, but at different rates for different soils. An exception is uniformly graded material, where some of the cement acts as a filler in the voids between the particles.

Cement modification Density

Minimal quantities of cement are sometimes used only to reduce the plasticity of unbound base and sub-base materials. This treatment is called 'cement modification'.

The density is almost as important as the cement content. On average, a 1 per cent increase in density leads to a 10 per cent increase in strength. The cement-stabilized material should be compacted immediately after mixing, because the cement begins to hydrate as soon as it comes into contact with water. If the compaction is delayed, some of the cemented bonds that have been formed will be broken down and lost.

Water

The presence of sufficient water in the fresh mixture is important both for the compaction and for the hydration reactions to proceed. The hydration is usually only affected at moisture contents which are too low to allow adequate compaction. Therefore, the moisture content should be chosen as close as possible to the optimum for the compaction equipment employed. If the optimum moisture content cannot be achieved, it is preferable to be on the wet side of the optimum rather than on the dry side. The stabilized layer should be kept continuously moist during the hardening process.

Temperature

The relation between strength and temperature for cement-stabilized materials is approximately linear, except for materials containing pozzolanic materials. Pozzolanic reactions are discussed later.

Cracks

Stabilized layers will crack due to shrinkage and changes in temperature and moisture content. The crack pattern is dependent on the strength of the material. Low strength materials will have frequent but narrow cracks. Cracks in capping layers and sub-base layers are not likely to cause any problems, as they are covered by a thick base. Cracks in a base, however, may be reflected through the wearing course and this may allow penetration of water down to the subgrade, leading to the deterioration of the pavement. In order to minimize the risk of reflection cracking, the strength of cement-stabilized materials is usually not allowed to exceed an upper limit. However, the most effective method to prevent reflection cracking is to cover the stabilized layer with a layer of unbound, granular material.

Mix design

The following procedure is recommended for selecting the cement content:

- 1 The optimum moisture content and the maximum dry density should be determined for the soil mixed with different amount of cement.
- 2 After mixing the soil, cement and water, the samples should be left two hours before being compacted using modified proctor compaction.
- 3 When the compaction tests have been completed, samples for strength tests should be mixed at the optimum moisture content obtained; again the mixtures should be left two hours before being compacted into cubes or test cylinders at 97 per cent of the maximum dry density.

- 4 The compacted samples should be cured in moist condition for seven days and soaked in water for seven days.
- 5 The samples should be crushed and an estimate made of the cement content needed to achieve the required strength.

Recommended strength requirements for different types of stabilized layers are listed in Table 13.5. As laboratory mixing is normally more efficient than field mixing, the cement content for field use is 1 or 2 per cent higher than that determined from laboratory tests.

13.5.3 Lime stabilization

Natural limestone ($\text{Ca}(\text{CO}_3)$) is used to adjust the pH of agricultural soils, but is not effective as a stabilizing agent in road pavements. The term 'lime' is used here to mean quicklime (calcium oxide CaO) and hydrated lime (calcium hydroxide $\text{Ca}(\text{OH})_2$). Quicklime is produced by heating limestone in a kiln until the carbon dioxide is driven out. Quicklime is usually 'slaked' with water, forming hydrated lime as a fine dry powder. It is not necessary to use high quality lime for soil stabilization. In many countries, low-grade limestone extracted locally and processed in temporary burning kilns is used in the production of lime for road construction.

Lime is effective as a stabilizing agent in most clayey soils, which are difficult to stabilize with cement. The stabilizing effect is due to the clay particles. Most clay particles are flat or elongated and, thus, have a large surface in relation to the unit weight. The clay particles have an ability to adsorb cations on to their surface. Moreover, most clay minerals are pozzolanic: in the presence of water, they will react chemically with lime to form cementing products.

Lime has an almost instantaneous effect on the plasticity of clayey soils. Even a small addition of lime will increase the plastic limit (PL), as shown in Figure 13.8. The effect on the LL is less pronounced. The overall result is a reduction in the PI and a more friable soil structure, which makes the soil easier to work and compact. The effect of lime on the plasticity is caused by the flocculation of the clay particles brought about by cation exchange in which cations, such as sodium and hydrogen, are replaced by calcium ions. The decrease in plasticity is followed by an increase in the strength of the soil, as measured by the CBR. If more lime has been added than needed for the cation exchange reactions, the lime will react chemically with the

Lime

Clay minerals

Reactions

Table 13.5 Recommended strength requirements for cement and lime stabilized sub-base and base materials

Type	Unconfined compressive cube strength (MPa)
Stabilized base (CB1)	3.0–6.0
Stabilized base (CB2)	1.5–3.0
Stabilized sub-base	0.75–1.5

Source: Overseas Road Note 31 (TRL 1993).

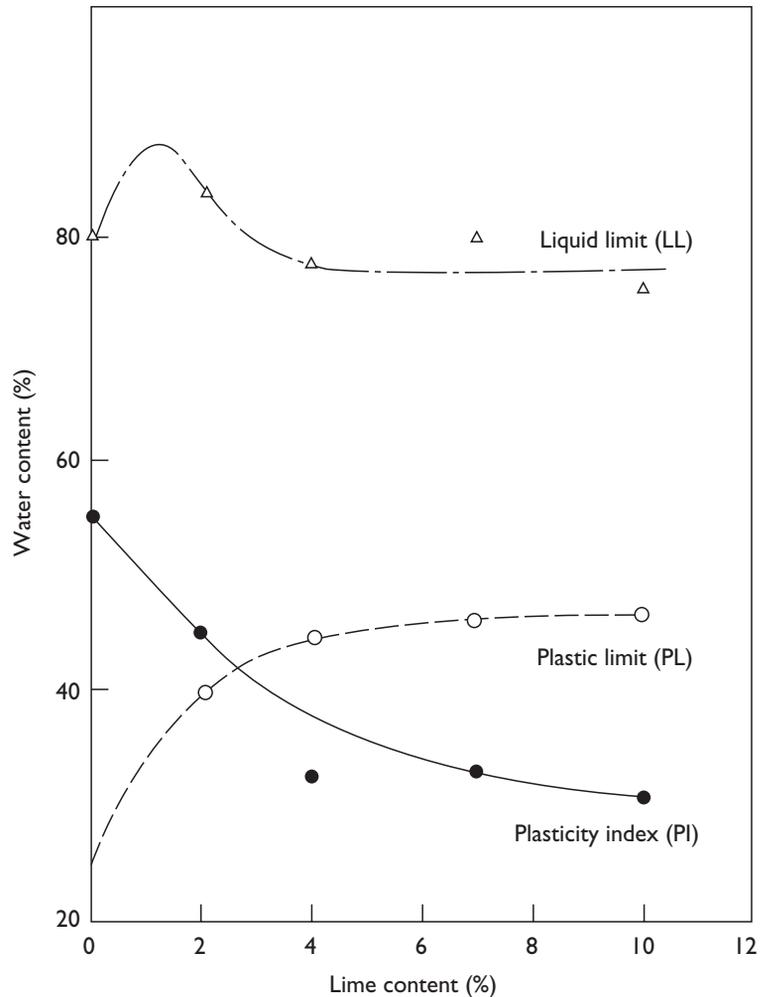


Figure 13.8 Effect of the addition of lime on London Clay.

Source: Sherwood (1993).

silicates and aluminates of most clay minerals and form cementing products similar to those in portland cement. In the presence of water, the cementing products will hydrate and, in time, will form a hard matrix similar to that in cement-stabilized materials.

Lime modification

Lime is often used only as a construction expedient to improve the workability of wet clayey soils without regard to any time-dependent increase in strength. Lime may also be used to reduce the plasticity of unbound base and sub-base materials. These treatments are usually called lime modification.

Lime type

The type of lime most widely used for soil stabilization is hydrated lime. Quicklime has a higher drying effect on wet soils, but quicklime is very caustic and creates the risk of skin and eye burns to site personnel. Even hydrated lime should be handled with care. Where the soil is dry and it is necessary to add water, the lime may be sprayed as a slurry in water, which reduces the dust problem.

Lime content

The strength of soil–lime mixtures increases with increasing lime content, but only up to a certain level. Nothing is gained by adding more lime than that corresponding to the content of reactive clay minerals in the soil.

A high density is equally important for both lime and cement-stabilized soils. However, because the pozzolanic reactions in soil–lime mixtures proceed much slower than the hydration of cement, any delay between mixing and compaction is far less critical for lime stabilized soils than for cement-treated materials. There may even be advantages in leaving the mixture for some hours after mixing to allow the lime to produce maximum effect on the plasticity. The material can then be remixed and compacted.

Density

Both lime and cement-stabilized materials should be compacted at optimum moisture content. The presence of sufficient water is essential both for compaction, and for the pozzolanic reactions and the hydration to proceed.

Water

At temperatures below 20–25°C, the pozzolanic reactions take place very slowly. Above 25–30°C, there is a large increase in strength for each degree of temperature increase. The rapid gain in strength with increasing temperature is an added reason why lime is more attractive than cement for soil stabilization in countries having a warm climate.

Temperature

Lime stabilized materials are subject to cracking for the same reasons as cement-stabilized materials, but the effects are less pronounced. Continuing pozzolanic reactions may result in ‘self healing’ of the cracks.

Cracks

If air has access to lime stabilized soils during the hardening process, the hydrated lime will react with carbon dioxide from the air and revert to calcium carbonate. The main distress caused by carbonation is loosening of the surface. To prevent carbonation, the stabilized layer should not be allowed to dry out during the hardening process, and the surface should be sealed as soon as possible.

Carbonation

The appropriate lime content depends on the pavement layer in which the lime treated soil will be used. Lime modified soils intended for use in capping layers should have a minimum CBR of 15 per cent. Lime stabilized soils in sub-bases and bases should satisfy the same strength requirements as cement-stabilized soils (as in Table 13.5). However, the curing period for lime stabilized test cubes or cylinders should be 21 days of moist curing, followed by seven days of soaking – in contrast to the seven-plus-seven days of curing used for cement-stabilized samples.

Mix design

Besides clay minerals, several ‘artificial’ materials have pozzolanic properties. The most important of these materials is pulverized fuel ash or fly ash, which is a by-product from power stations burning coal-dust. Mixtures of lime and fly ash may be used for stabilizing non-clayey granular materials. Another use of fly ash is as an additive to cement. When cement reacts with water, hydrated lime is produced in addition to the strength-giving hydrated calcium silicates and aluminates. If fly ash is present, it will react with this lime to produce further cementing materials. ‘Fly ash cement’ may contain up to 30 per cent fly ash. The ash produced from burning agricultural wastes may be rich in reactive silica. Rice husk ash, which is produced in great quantities throughout Asia, has proved to be an excellent pozzolan. The ash from burning of bagasse (crushed sugar cane) is also known to have pozzolanic properties.

Pozzolans

13.5.4 Construction of stabilized layers

Stabilized layers may be constructed using either the mix-in-place method or the plant-mix-method. Mix-in-place operations take place on the construction site, where the stabilizer is mixed with the in-situ subgrade soil, or with borrow material

Options

placed on the formation of the road. In plant-mix operations, the stabilizer and the soil are mixed at a stationary mixing plant and then transported to the road site.

Mix-in-place

In the mix-in-place method, the material to be stabilized is first shaped to the required level and crown. An initial scarification or pulverization is then performed to a specified depth and width. After the soil has been loosened, the stabilizer is spread either by mechanical spreaders or by hand. Mechanical spreaders give the most uniform distribution. Attention should be paid to the dust problem. When spreading stabilizer by hand, the labourers should be provided with protective clothing and face masks. Lime is caustic when coming into contact with moisture. Therefore, it is particularly important that labourers do not get lime in their eyes and respiratory passages. After spreading, the stabilizer is mixed with the soil using special travelling mixers (rotavators). Motor graders are sometimes used, but they are inefficient for pulverizing clayey soils and too slow for processing cement-stabilized soils. Agricultural disc ploughs or harrows may be used for lime modification of overly wet soils. Hand mixing has also been used, but this highly inefficient method should be discouraged. If it is necessary to add water to obtain the required optimum moisture content, this may be done through a sprayer in the mixer. In case the mixer is not equipped with a spray system, water should be added and mixed with the soil prior to spreading of the stabilizer. When stabilizing dry soils with lime, the lime is most conveniently distributed as a slurry in water.

Plant-mix

Plant-mix is produced in a central plant and the mixture hauled to the job site in trucks. Plant mixing is used principally for cement stabilization of gravelly materials, since clayey materials cannot usually be mixed using this process. Plant mixing enables good control on mix proportions, but the output is lower than in the mix-in-place process. On the construction site, the mixture is spread by an asphalt spreader or a spreader box. If graders are used, it is difficult to obtain the right levels and thickness of the stabilized layer, and many of the advantages of plant-mixing are lost.

Compaction

The stabilized layer must be compacted as soon as possible after mixing. Most specifications require compaction of cement-stabilized soils to be completed within two hours. Delays in lime stabilization are less critical, but the compaction should still be finished quickly, particularly in hot climates where evaporation of water and carbonation of lime may create problems. The compaction may be delayed only when lime modification is used to provide a more workable soil.

Curing

The newly finished layer must be properly cured: loss of water must be prevented for the first seven days or more. Curing is necessary to

- ensure continuous hydration of the stabilizer;
- reduce shrinkage;
- reduce the risk of carbonation.

Good curing is particularly important, but also very difficult in hot climates. Frequent spraying of water is commonly used, but this method is likely to leach stabilizer from the top of the layer. Also, if the spraying is interrupted and the surface allowed to dry, even for a short while, the curing will be ineffective. Spraying of water can be more effective if a layer of 30–40mm sand is first placed on the surface of the stabilized layer. The sand is removed when the curing is complete. The most effective curing method, however, is to apply a very light spray of water

followed by a viscous cutback bitumen or a slow setting emulsion. This will serve as a membrane preventing evaporation of water from the stabilized layer.

References

- Millard, R.S. (1993) 'Road building in the tropics', *TRL State of the Art Review No. 9*, London: H.M. Stationery Office.
- Parsons, A.W. (1992) 'Compaction of soils and granular materials', *TRL State of the Art Review*, London: H.M. Stationery Office.
- Sherwood, P.T. (1993) 'Soil stabilization with cement and lime', *TRL State of the Art Review*, London: H.M. Stationery Office.
- TRL (1993) 'A guide to the structural design of bitumen-surfaced roads in tropical and sub-tropical countries', *Overseas Road Note 31*, 4th edition, Crowthorne: Transport Research Laboratory.
- TRL (1997) 'Principles of low cost road engineering in mountainous regions, with special reference to the Nepal Himalaya', *Overseas Road Note 16*, Crowthorne: Transport Research Laboratory.

Asphalt pavement materials

Richard Koole and Bent Thagesen

14.1 Asphalt pavements

Asphalt pavements consist of selected mineral aggregates bound together by a bituminous binder. Asphalt includes a multitude of different pavement types, ranging from thin surface dressings to thick layers of asphalt concrete. This chapter deals with bituminous binders and the most widespread types of asphalt pavement materials.

14.2 Bituminous binders

14.2.1 Bitumen

Definition

Bitumen (called ‘asphalt cement’ in the United States) is a black to dark brown sticky material, composed principally of high molecular-weight hydrocarbons. In some countries, it can be found as a component of natural rock asphalt, but most bitumen is derived from the distillation of crude oil. Not all crude oils are sufficiently rich in heavy components to yield bitumen economically. Crude oils from Indonesia and Nigeria are examples of so-called light crudes containing very little suitable heavy residue. Heavy crude oils from the Middle East and South America usually have a large content of heavy residue suitable for bitumen (Millard 1993).

Characteristics of bitumen

Bitumen is a thermoplastic material that gradually softens, and eventually liquefies when heated. Bitumen is characterized by its consistency at certain temperatures. The consistency is measured by a penetration test, a softening point test and a viscosity test. More recently developed characterization methods use a ‘complex modulus’ at different temperatures.

14.2.2 Bitumen tests

Penetration

Penetration is determined using an empirical test method. A sample of the bitumen in a small container is stabilized in a water bath at a standard temperature (normally 25°C). A prescribed needle, weighing 100 g, is allowed to bear on the surface of the bitumen for 5 sec. The penetration is defined as the distance, in units of 0.1 mm, that the needle penetrates into the bitumen. The test is illustrated in Figure 14.1.

Softening point

Determination of the softening point is also an empirical test. A hot sample of bitumen is poured into a brass ring and allowed to cool. The top is then levelled with a hot spatula, and the ring placed in a holder and lowered into a beaker with water

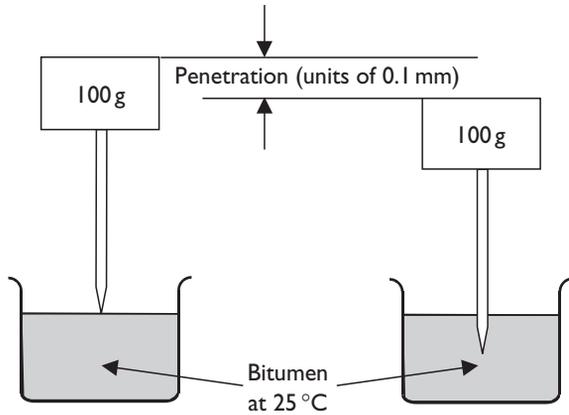


Figure 14.1 Penetration test.

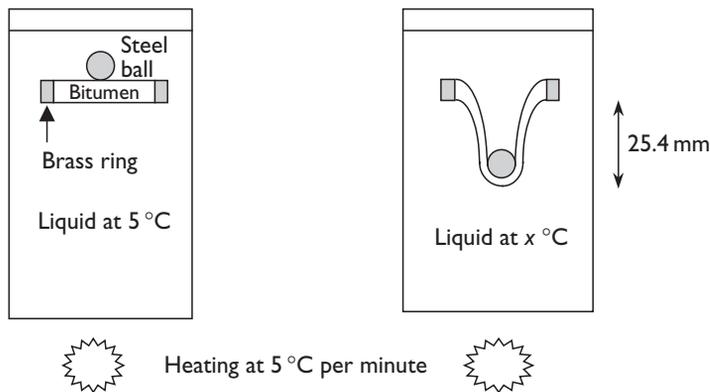


Figure 14.2 Softening point test.

or glycerine at 5°C. It is then loaded with a steel ball of specified weight and diameter. The water (or glycerine) is then heated at a rate of 5°C per minute. As the water is heated, the disc of bitumen in the ring will sag under the weight of the ball. When the disc has sagged over a distance of 25.4 mm, the temperature of the water at that instance is recorded as the ‘softening point’ of the bitumen. The softening point is often also called the ‘ring and ball’ temperature and, for typical bitumens used for road pavements, is roughly in the range of 45–60°C. The penetration value at 25°C and the softening point are both measures of consistency of the bitumen at different temperatures. When considered together, they are a measure of temperature susceptibility of the bitumen. The softening point test is illustrated in Figure 14.2.

Viscosity is measured at 60°C using a capillary viscometer. As bitumen is too viscous to flow readily at 60°C, a prescribed vacuum is applied to the viscometer to induce flow. The time in seconds required for the bitumen to flow between two timing marks is measured. The recorded time multiplied by a calibration factor for the particular viscometer gives the viscosity in poise. Note that, since the penetration test is empirical and the viscosity test is scientific, there is no direct relationship between the two.

Viscosity

Penetration grades	Bitumen is available commercially in several standard grades. In many countries, the grades are based on the penetration value. The grade is usually expressed in a penetration value bracket, for example, 80–100. The British Standard (BS) specifies 10 different grades ranging from pen 15 to pen 450. Earlier standards in the United States specified five types, with pen 40–50 as the hardest and pen 200–300 as the softest.
Viscosity grades	Traditionally, paving bitumens have been specified in terms of their penetration, but the measurement of viscosity provides a more accurate method of specifying binder consistency and a more effective method of determining the temperature susceptibility of the bitumen. American Association of State Highway Officials (AASHTO) has two series of viscosity grade bitumens. One is denoted AC (asphalt cement), followed by a number indicating the viscosity in hundreds of poises at 60 °C. The second series is denoted AR (aged residue) followed by a number indicating the viscosity in poises (not hundreds of poises) at 60 °C after the bitumen has been aged. The ageing is obtained by exposing a film of bitumen in a revolving jar to a jet of hot air for a prescribed period of time. The procedure is intended to subject the sample to hardening conditions approximating to those that occur in a hot-mix asphalt plant, and is known as the ‘rolling thin film oven test’ (RTFOT).
Performance grading	A new grading system was introduced in the 1990s as a result of the US ‘Strategic Highway Research Program’. It is called ‘performance grading’ (PG). In principle, it stipulates a temperature window within which the bitumen meets certain criteria. For instance, a PG 64–16 bitumen meets these criteria at temperatures between +64 °C and –16 °C, while a PG 82–28 will meet them between +82 °C and –28 °C. The wider the temperature window, the more stringent the requirements on the bitumen. The PG grading required in any one region will depend on the climate and the seasonal temperature variations, as well as the intensity of the traffic.
Liquefied bitumen	Bitumen must be liquefied before it can be sprayed on a surface or mixed with aggregate. Heating is one way of making the bitumen sufficiently fluid. Other ways are to dissolve the bitumen in a solvent, or to disperse the bitumen in water as an emulsion. The two products are known as ‘cutback bitumen’ and ‘emulsified bitumen’, respectively. When cutback bitumen is used for pavement construction, the solvent will evaporate on completion of the work, leaving the bitumen to perform its function as a binder. With emulsified bitumen the emulsion will ‘break’, with the water evaporating off, and precipitating a bitumen film on the mineral aggregate.
14.2.3 Cutback bitumen	
Properties	Cutback bitumen is produced with different rates of curing (hardening) and different degrees of viscosity. The rate of curing is controlled by the volatility of the solvent used. The viscosity depends principally on the proportion of solvent to bitumen. The more viscous grades require heating to make them fluid enough for use. Cutbacks are popular because of their ease of handling and application – they tend to be ‘forgiving’ of poor workmanship. However, their use is decreasing because of safety concerns (flammability) and for environmental reasons related to the evaporation of hydrocarbons into the atmosphere. Their role is being taken over by emulsions.
AASHTO groups	AASHTO specifies three groups of cutback bitumen: rapid curing (RC) where petrol/gasoline is used as a solvent, medium curing (MC) containing kerosene and

slow curing (SC) made with diesel oil. The viscosity is measured with a capillary tube viscometer at 60°C. The flow through the viscometer is induced by the gravity, and vacuum is not needed. The viscosity is expressed in centistokes whereas, for pure bitumen, the viscosity is expressed in poise. The units of poise and stokes are related to each other through the density of the tested material.

The British Standard provides specifications for three viscosity grades of cutback bitumen intended for use in surface dressings. Measurement is with a discharge viscometer, and viscosity is defined as the time in seconds for 50 millilitres of the binder to flow through a standard orifice at 40°C.

BS groups

14.2.4 Emulsified bitumen

In the emulsification process, hot bitumen is divided into minute globules and dispersed in water. The machine used in this process is called a colloid mill. An emulsifying agent is added to the water to assist the emulsification, and discourage coagulation of the bitumen globules. The emulsifying agent provides the bitumen globules with electrical surface charges that prevent the small particles from coalescing. If the charges are negative, the emulsion is called 'anionic'. If the charges are positive, the emulsion is called 'cationic'. Emulsions may require a small amount of heating to make them fluid enough for use.

Emulsification process

When emulsified bitumen is sprayed onto a road surface or mixed with aggregate, the bitumen separates from the water. This is either through evaporation of the water (anionic emulsions) and/or neutralization of the electric charges (cationic emulsions). The manner and rate at which the emulsion breaks, or sets, depends largely on the properties of the emulsifying agent and the relative proportions of bitumen and water. Both anionic and cationic emulsions are manufactured in several grades. BS distinguishes between 'rapid', 'medium' and 'slow breaking' types of emulsified bitumen. AASHTO uses the terms 'rapid', 'medium' and 'slow setting'. Rapid and medium setting emulsions are generally of a cationic nature. They are used in surface dressings because the contact with the stone initiates breaking and shortens the time for the road to be opened to traffic. Slow setting emulsions are generally anionic, and are used in asphalt mixes as they allow more homogeneous mixing with the aggregate before setting.

Setting of emulsion

14.2.5 Modified bitumen

During the last decade different additives have been developed to improve the properties of bitumen. Most additives are based on rubber or polymers. Rubberized bitumen requires mixing at very high temperatures and, because of this, is not favoured in many countries.

Additives

The addition of polymers can make the bitumen more resistant to loading and/or less susceptible to temperature variations. In addition, some polymers will improve adhesion of the bitumen to the stones, particularly in the wet, and improve the resistance to cracking. However, not all polymers and all bitumens are compatible, in the sense that they will form a homogeneous dispersion and are stable during storage. Polymer-modified bitumens are considerably more expensive than conventional bitumens. Their use is only justified where traffic loading is heavy, or where other

Polymer-modification

severe conditions exist, and where their use results in lower road maintenance requirements.

14.3 Surface dressing

14.3.1 Characteristics

Definition	A surface dressing (surface treatment in the United States) is a wearing course made by a thin film of binder sprayed onto the road surface and immediately covered with a layer of stone chippings of uniform size. The thin film of binder acts as a waterproofing seal preventing the entry of surface water into road base. The chippings protect the film of binder from damage by traffic, and form a skid-resistant and dust-free wearing surface. The process may be repeated to provide double or triple layers of chippings.
Use	Surface dressing is an effective maintenance technique that is capable of extending the life of a structurally sound road pavement. Surface dressing can also provide an effective and inexpensive surface for new pavements with a traffic flow of up to 500 vehicles per lane per day. When used as a maintenance operation to an existing asphalt pavement, a single surface dressing is normally adequate. For new road bases, a double surface dressing should be applied. A well-designed and constructed surface dressing should last five years or more before resealing becomes necessary.
Design guide	A guide to surface dressing has been published as <i>Overseas Road Note 3</i> by the Transport Research Laboratory (TRL and DFID 2000). The main points from this are presented here.

14.3.2 Selection of chippings

Strength	Ideally, the chippings should be single-sized, cubical, strong, durable, clean and dust-free. Specifications for maximum aggregate crushing value (ACV) typically lie in the range of 20–35. In wet climates, the chippings should be resistant to polishing under the action of traffic. British specifications call for a minimum polished stone value (PSV) of 45–60, depending on traffic and site characteristics. In practice, the chippings available often fall short of the ideal.
Size	The action of the traffic on a surface dressing gradually forces the chippings down into the underlying surface. This embedment process occurs more rapidly when the underlying road surface is soft and when the volume of commercial traffic is high. Consequently, large chippings are required on soft surfaces or where traffic is heavy, and small chippings should be used for hard surfaces and light traffic. The nominal size of chippings used for surface dressings is usually 6, 10, 14 or 20 mm. The appropriate size is selected from Table 14.1.
Double surface dressings	For double surface dressings, the size of chippings for the first layer should also be selected according to Table 14.1. The chippings for the second layer should preferably have an average least dimension (ALD – see later) of not more than half of the chippings used in the first layer to promote the formation of a mosaic, with good interlock between layers. Sand may sometimes be used as an alternative to chippings for the second layer. However, if the existing surface is very hard, such as a newly

Table 14.1 Recommended nominal chipping size in mm for surface dressing

Type of surface	Approximate number of commercial vehicles with an unladen weight greater than 1.5 tonnes carried per day in the design lane				
	2,000–4,000	1,000–2,000	200–1,000	20–200	Less than 20
Very hard	10	10	6	6	6
Hard	14	14	10	6	6
Normal	20 ^a	14	10	10	6
Soft	*	20 ^a	14	14	10
Very soft	*	*	20 ^a	14	10

Source: Overseas Road Note 3 (TRL and DFID 2000).

Notes

The size of chipping specified is related to the mid-point of each lane traffic category; lighter traffic conditions may make the next smaller size of stone more appropriate.

* Unsuitable for surface dressing.

a Very particular care should be taken when using 20mm chippings to ensure that no loose chippings remain on the surface when the road is open to unrestricted traffic, as there is a high risk of wind-screen breakage.

constructed cement stabilized base, a 'pad coat' of 6-mm chippings should be applied first. This is then followed by a second layer with 10- or 14-mm chippings.

The ALD of chippings is the average thickness of a single layer of chippings when they have bedded down into their final interlocked positions. It is assumed that the particles will settle with their shortest side vertical. ALD of the selected chippings may be measured manually on a representative sample.

ALD

14.3.3 Selection of binder

The correct choice of binder is critical. The binder must be sufficiently fluid at road temperature to wet the road surface and the chippings. At the same time, the binder must be sufficiently viscous not to drain off from the road surface, and strong enough to retain the chippings when the road is opened to traffic. Both pure bitumen, cutback bitumen and emulsified bitumen may be used as binder for surface dressings. As noted earlier, the use of cutbacks is decreasing. Emulsions have a great potential in developing and emerging countries, because they can be manufactured locally using small mobile plants. Also, emulsions can be made with a harder bitumen than cutbacks, and a hard bitumen is advantageous in hot climates. The use of emulsified bitumen for surface dressings is generally increasing.

Requirements

Figure 14.3 shows the permissible range of binder viscosity at various road temperatures, and the appropriate grades of penetration grade bitumens and cutbacks to be used.

Bitumen and cutback

Cationic emulsion with a bitumen content of 70–75 per cent is recommended for most surface dressing work. Most emulsified bitumens have a low viscosity and easily wet the road surface and chippings. Because of the low viscosity, emulsion may not be appropriate if a high rate of spray is required, as in the case for large-sized chippings. The light-fluid emulsion may drain off from the crown of the road before breaking.

Emulsion

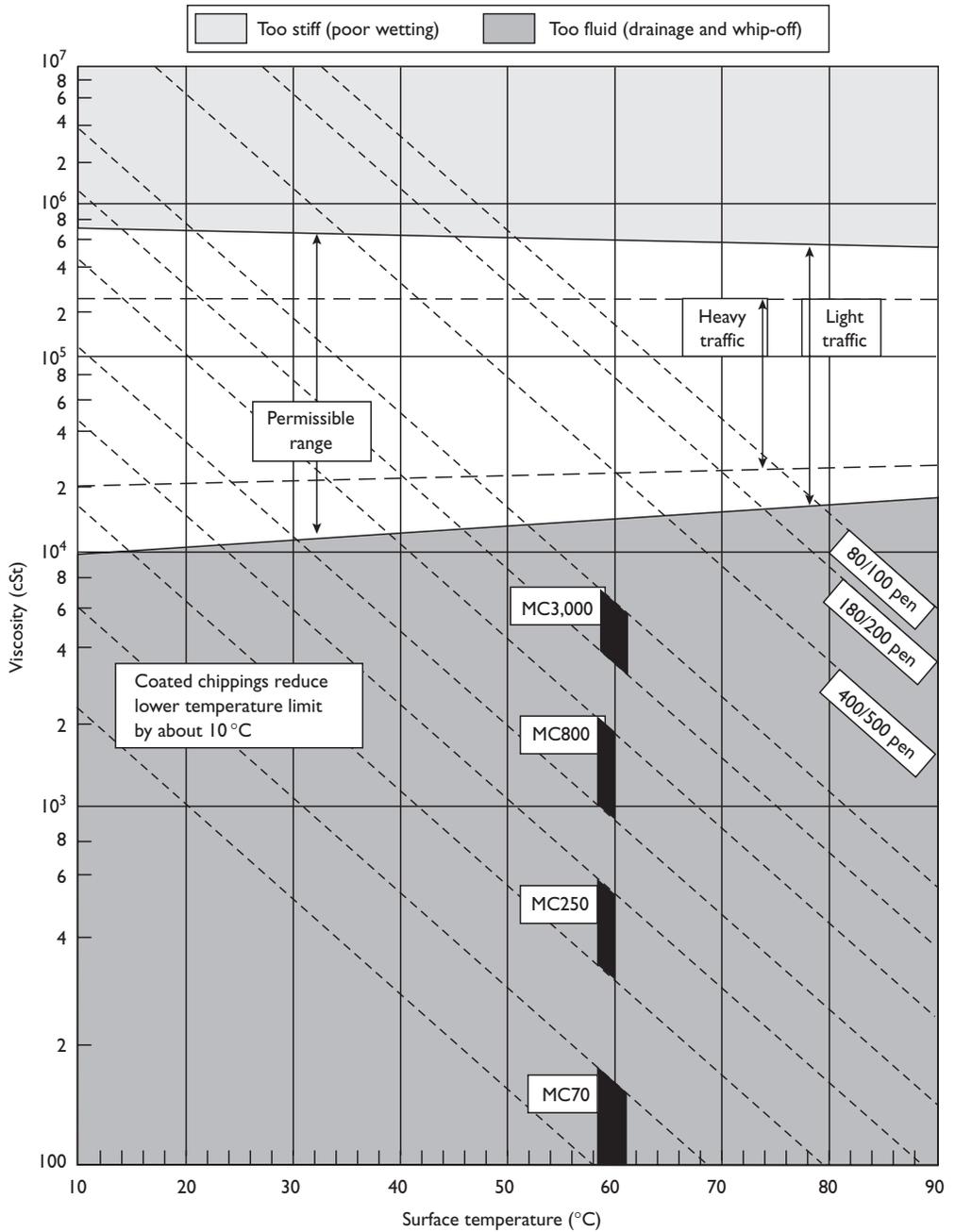


Figure 14.3 Choice of binder for surface dressing.
 Source: Overseas Road Note 3 (TRL and DFID 2000).

14.3.4 Rate of spread of binder and chippings

Binder

The rate of application of binder depends on the type of chippings, the level of traffic, the condition of the existing road surface, and the climate. An appropriate factor for each of the four sets of conditions is selected from Table 14.2 and added together to give a total weighing factor. The ALD and the weighing factor are then

Table 14.2 Condition factors for determining the rate of application of binder for surface dressing

Vehicles per lane per day		Existing surface	
0–50	+3	Untreated or primed base	+6
50–250	+1	Very lean bitumen	+4
250–500	0	Lean bitumen	0
500–1,500	–1	Average bitumen	–1
1,500–3,000	–3	Very rich bitumen	–3
>3,000	–5		
Climate		Type of chippings	
Wet and cold	+2	Round/dusty	+2
Tropical (wet and hot)	+1	Cubical	0
Temperate	0	Flaky	–2
Semi-arid (dry and hot)	–1	Pre-coated	–2
Arid (very dry and hot)	–2		

Source: Overseas Road Note 3 (TRL and DFID 2000).

used with Figure 14.4 to obtain the required rate of binder application. The rate of spread of cutback bitumen MC/RC 3,000 is determined by the intercept between the factor line and the appropriate ALD line. In the rare cases when cutbacks with lower viscosity are used, the rate of spread should be increased to allow for the additional percentage of solvent used. For penetration grade bitumen, the rate determined from the chart should be reduced by 10 per cent. For emulsified bitumen, the rate should be multiplied by $(90/b)$, where b is the percentage of bitumen in the emulsion. For slow traffic or steep climbing grades, the basic rate of spread should be reduced by approximately 10 per cent. For fast traffic or steep down grades, the rate should be increased by approximately 10 per cent.

Chippings should be spread at a rate corresponding to a single, tightly packed layer, plus a 10 per cent allowance to ensure complete coverage. An estimate of the chipping application rate, assuming that the chippings have a loose density of $1,350 \text{ kg/m}^3$, can be obtained from the following equation:

$$\text{Chipping application rate (kg/m}^3\text{)} = 1.364 \text{ ALD}$$

14.3.5 Construction of surface dressing

Before applying a surface dressing, the surface must be carefully prepared. Existing asphalt pavements must be repaired and all pot-holes patched with asphalt materials to produce a surface texture similar to the adjacent pavement. Traffic should be allowed to run over the patched road for some months. The surface must be dust-free and dry so, immediately prior to spraying the binder, the road should be swept clean. The surface dressing operation should never be started when rain is expected within the first few hours.

The binder must be applied uniformly at the correct rate of spread. Watering cans may be used for minor works such as patching. A more controllable method is to use

Chippings

Preparation

Spreading of binder

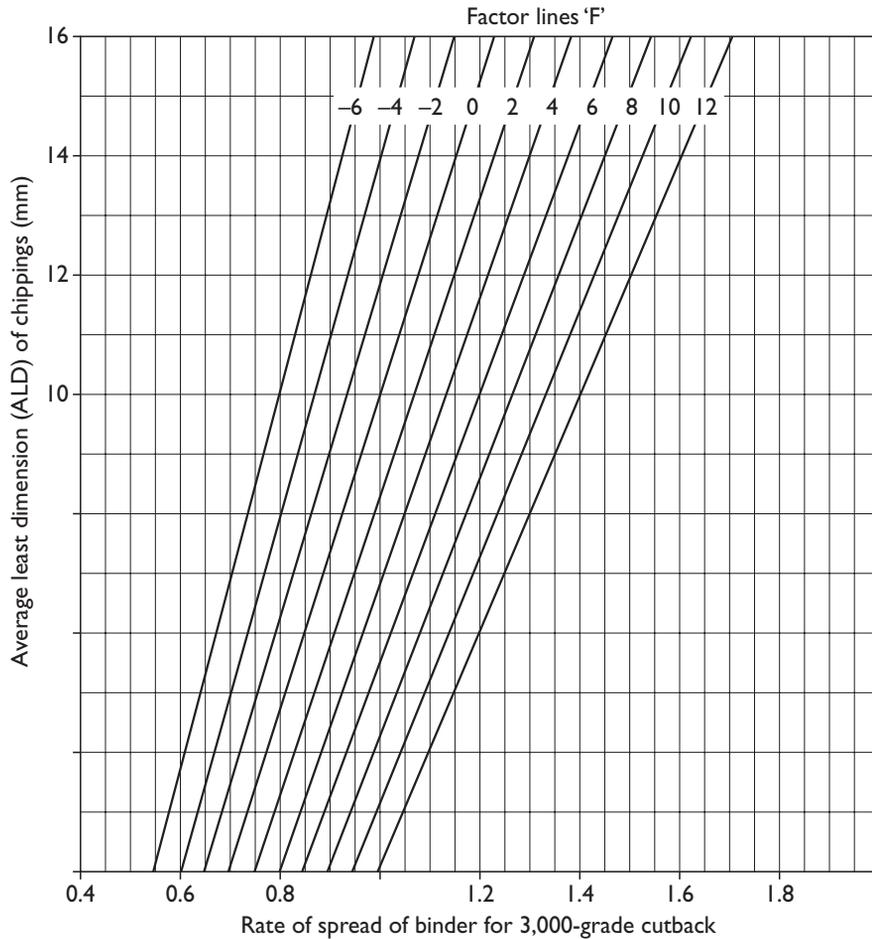


Figure 14.4 Rate of spread of binder for 3,000-grade cutback (kg/m²).

Source: *Overseas Road Note 3* (TRL and DFID 2000).

hand lances, since these can produce an acceptable, uniform rate of spread. However, it is difficult to achieve a specified rate of spread with either method. Successful spreading of binder on large areas requires the use of a mechanical bulk distributor. Distributors spray the bitumen through a spray bar to which the binder is delivered by a pump, or under pressure from a heated insulated storage tank. There are basically two types of spray jets, slotted jets and whirling spray jets. Suitable spraying temperatures for penetration grade and cutback bitumens are given in Table 14.3. Emulsions with a bitumen content of 70–75 per cent can be applied through whirling spray jets at a temperature between 70 and 85 °C.

Spreading of chippings

Immediately after the binder has been applied, the chippings should be spread. This ensures maximum possible wetting of the chippings. The chippings can be spread by hand with a good result, but a mechanical chip spreader facilitates rapid application and an even distribution.

Rolling

Following the distribution of chippings, the layer should be rolled to seat the aggregate firmly in the bitumen. Traditionally, steel wheeled-rollers have been used

Table 14.3 Spraying temperatures for binders

	Whirling spray jets		Slotted jets	
	Min °C	Max °C	Min °C	Max °C
<i>Cutback grades</i>				
MC30	50	60	40	50
RC/MC70	65	80	55	70
RC/MC250	95	115	80	90
RC/MC800	115	135	105	115
RC/MC3,000	135	150	120	13
<i>Penetration grades</i>				
400/500	160	170	140	150
280/320	165	175	150	160
180/200	170	190	155	165
80/100	180	200	165	175

Source: Overseas Road Note 3 (TRL and DFID 2000).

for this purpose, but they tend to crush weak aggregate. Pneumatic-tyred rollers are much preferred.

As it is necessary to apply a slight excess of chippings, some loose material will inevitably be left on the new surface dressing. This excess may be removed by brooming once the traffic has been permitted to run over the road surface for a few days.

Brooming

Construction procedures for double surface dressings are essentially the same as those for single surface dressings. Traffic should be allowed to run on the first dressing for two to three weeks or more before the second dressing is applied. If the surface is contaminated with dirt during this period it must be thoroughly swept before the second dressing is applied.

Double surface dressing

14.4 Premixed asphalt

14.4.1 Types of premixed asphalt

Premixed asphalt is a paving material manufactured by mixing aggregates, filler and bitumen. Premixed asphalt is used in construction of wearing courses binder courses and bases. Most premixed asphalt is mixed and placed hot, hence the name 'hot-mix'. This type of asphalt can be divided into mixes where traffic stresses are transmitted mainly through the aggregate structure, and mixes where the stresses are transmitted mainly through the mortar consisting of bitumen, filler and fines. Asphalt concrete, by far the most common type of premixed asphalt, is of the first type. The aggregate in asphalt concrete is continuously graded. Bitumen macadam is similar to asphalt concrete, but has a less dense aggregate structure. Stone mastic asphalt and drainage asphalt are also mixes of this type, but use gap-graded aggregates. Hot-rolled asphalt is in the second category. The aggregate is gap-graded. The content of bitumen/filler/fines mortar is relatively high and determines the performance of the mix. Stresses are transmitted through the mortar.

14.4.2 Hot-mix asphalt

Recommendations for materials intended for use in hot-mix asphalt in tropical and sub-tropical countries are given in *Overseas Road Note 19*. The recommendations are outlined here.

Objectives

To facilitate placing and compaction of the asphalt, the fresh mix should have good workability. In order to perform satisfactorily, the finished asphalt layer needs to have

- high stiffness to reduce the stresses transmitted to the underlying layer;
- high resistance to deformation;
- high resistance to fatigue;
- high resistance to weathering (good durability);
- low permeability to prevent ingress of air and water;
- sufficient surface texture to provide good skid resistance in wet weather (only wearing courses).

Some of these requirements are conflicting; for example, good durability and high resistance to deformation. The key to asphalt-mix design is to produce a mix that possesses an acceptable balance of properties. Mix requirements are more critical in countries with extreme climates, because of higher temperature ranges and, often, higher axle loads.

Durability

The durability of a mix is associated with the resistance to age hardening of the bitumen. One way of reducing age hardening is to increase the bitumen content. This reduces the content of air voids in the mix and increases the thickness of the bitumen film coating the aggregate particles. This is why specifications for asphalt concrete normally require that a minimum amount of the void-space between the aggregate particles of the compacted mix should be filled with bitumen.

Resistance to deformation

However, increasing bitumen content reduces the stiffness and the resistance to plastic deformation of the asphalt. Plastic deformation is the most serious form of failure for asphalt pavements because the affected layer must be removed before the pavement can be rehabilitated. To avoid plastic deformation in asphalt concrete, a minimum of 3 per cent of the bulk volume of the compacted mix should be filled with air. If the total air volume decreases to less than this, stress transfer through stone to stone contact switches to the bitumen-fines mortar, and plastic deformations occur.

14.4.3 Materials for hot-mix asphalt

Aggregates

Aggregates are usually categorized as ‘coarse aggregate’ and ‘fine aggregate’:

- Coarse aggregate is the mineral aggregate retained on the 2.36-mm sieve; it is produced by crushing rock or natural gravel to obtain angular, rough-textured particles with good mechanical interlock.
- Fine aggregate is material passing the 2.36-mm sieve; it can be crushed rock or natural sand.

The aggregate should have the following characteristics:

- Be clean
- Be low-absorptive

- Be angular in shape
- Have good strength
- Be resistant to polishing (wearing course only)
- Be weather-resistant (sound)
- Have good affinity with bitumen.

Table 14.4 shows recommended requirements for aggregate.

Filler comprises the fines passing the 75- μm sieve. The coarse and fine aggregates contain filler in the form of naturally fine material. The natural filler collected in the filters or cyclones ('baghouse fines') may not have desirable properties, or may be in excess of requirements. Crushed rock fines, portland cement or hydrated lime may be added to obtain the required proportion and quality of filler in the mix. Addition of cement or hydrated lime is particularly useful as it improves the adhesion of the bitumen to the aggregate. It acts as an 'anti-stripping' agent.

Specifications for bitumen most appropriate for use in tropical countries are given in *ORN19*. Monitoring of bitumen quality requires relatively sophisticated laboratory equipment and, other than testing consistency, many asphalt producers rely on the test certificates presented by the bitumen manufactures (oil refineries).

Filler

Bitumen

14.4.4 Mix design of asphalt concrete

Combined aggregates for asphalt concrete should be well graded (continuously graded). Table 14.5 shows the recommended grading limits.

The bitumen used for production of asphalt concrete, where high road surface temperatures prevail, is usually pen 40–50 or pen 60–70. The bitumen content depends on the aggregate and the expected traffic. Normally, it lies between 4 and 7 per cent by mass of total asphalt mix. The exact 'design bitumen content' is determined by use of a mix design method.

Mix design for asphalt concrete is commonly carried out using the 'marshall' method. In the 1990s, the US Strategic Highway Research Program developed a new comprehensive design procedure called the 'Superpave' method. However, this method involves detailed testing of bitumen and the use of a 'gyratory' compactor

Aggregate

Bitumen

Design method

Table 14.4 Requirements for aggregates for hot-mix asphalt

Property	Test	Percentage
Cleanliness	Sand equivalent	>35–40 depending on ESA
Absorption	Water absorption	<2
Particle shape	Flakiness index	<35
Strength	Aggregate crushing value	<25
	Los Angeles abrasion	<30 for wearing course <35 for binder course
Polishing	Polish stone value	<45–70 depending on traffic
Soundness	Sodium sulphate test	<10 for coarse aggregate
		<16 for fine aggregate
Bitumen affinity	Coating retained after immersion	>95

Adapted from: *Overseas Road Note 19* (TRL and DFID 2002).

Table 14.5 Recommended particle size distribution for combined aggregate for use in asphalt concrete

Sieve size (mm)	Passing sieve (percentage by mass)				
	Wearing course			Binder course	Base
	Nominal maximum particle size (mm)			25	37.5
	9.5	12.5	19		
50					100
37.5				100	90–100
25			100	90–100	—
19		100	90–100	—	56–80
12.5	100	90–100	—	56–80	—
9.5	90–100	—	56–80	—	—
4.75	55–85	44–74	35–65	29–59	23–53
2.36	32–67	28–58	23–49	19–45	15–41
0.3	7–23	5–21	5–19	5–17	4–16
0.075	2–10	2–10	2–8	1–7	0–6

Source: *Overseas Road Note 19* (TRL and DFID 2002).

for mix design. Road administrations in developing and emerging countries will often encounter difficulties with complex test methods, and the marshall method will probably remain in use in many countries for a number of years. However, it should be noted that the marshall method cannot be used to design asphalt mixes with aggregate particles larger than 25 mm. Design of asphalt concrete binder courses and bases normally rely on empirical knowledge. The design procedure now described applies only to asphalt concrete wearing courses.

The marshall method

The marshall method uses impact compaction with a ‘marshall’ hammer to produce cylindrical test specimen (diameter 102 mm and height about 64 mm). A standard procedure for heating and mixing the selected aggregate and bitumen is used. First the design bitumen content is estimated based on experience. Then test specimens are produced at the estimated design bitumen content and at two increments of 0.5 per cent above and below the estimate. In order to ensure provision of adequate data, three specimens are prepared at each of the five different bitumen contents. The specimens are tested to determine their volumetric composition and their strength characteristics.

Density test

First, the specimens are subjected to a density test. This test consists of weighing the specimen in air and in water. In order to determine the volume of the specimen, including open voids in the surface, the mass is also determined after the specimen has been immersed in water and its surface blotted with a damp towel to dry the surface without removing the water in the surface voids. The ‘marshall density’ or bulk specific gravity (BSG) of the specimen is calculated using the formula:

$$\text{BSG} = \frac{M_d}{M_{sd} - M_i} \gamma_{wt}/m^3$$

where M_d is the mass of dry specimen; M_i the mass of immersed specimen; M_{sd} the mass of surface-dry specimen; and γ_w the specific gravity of water.

The volume of aggregate (V_a) and the volume of bitumen (V_b) are calculated using:

$$V_a = \frac{(100 - b)}{\gamma_a} \text{BSG \%}$$

$$V_b = \frac{b}{\gamma_b} \text{BSG \%}$$

where b is the mass of bitumen in per cent of mass of mix; γ_a the apparent specific gravity of aggregate; and γ_b the specific gravity of bitumen.

The next step is to calculate the voids in aggregate (VMA), the air voids in mix (VIM) and the voids in aggregate filled with bitumen (VFB) using:

$$\text{VMA} = 100 - V_a \%$$

$$\text{VIM} = 100 - V_a - V_b \%$$

$$\text{VFB} = \frac{V_b}{100 - V_a} 100 \%$$

Finally, each cylindrical specimen is subjected to a stability and flow test in a marshall testing machine. After being heated to 60°C in a water bath, the specimen is placed in the marshall machine between two collar-like testing heads, and compressed radially at a constant rate of displacement (see Figure 14.5). The maximum load resistance (in newtons) is recorded as the ‘marshall stability’. The corresponding total deformation in millimetres of the specimen is recorded as the ‘flow’.

When testing is complete, plots are prepared, as shown in Figure 14.6, for bitumen content versus:

- Bulk specific gravity (BSG)
- Voids in aggregate (VMA)
- Air voids in mix (VIM)
- Voids in aggregate filled with bitumen (VFB)
- Stability
- Flow.

ORNI9 recommends that the bitumen content giving 4 per cent air voids is used as a starting point for selecting the design asphalt content (5 per cent if $\text{ESA} > 5 \times 10^6$). At this bitumen content, the voids in aggregate, voids in aggregate filled with bitumen, stability and flow are determined by interpolation from the graphs. These properties are then compared to the design criteria in Table 14.6. The final design asphalt content is chosen as a compromise to balance all the mix properties. It should be ensured that the mix remains within specifications for the chosen design bitumen content ± 0.3 per cent (normal production tolerance). It is also important that the design bitumen content is chosen to be slightly less than that which gives the minimum voids in aggregate. More bitumen than that corresponding to the minimum voids in aggregate causes an increase of these voids. This indicates that the aggregate

Volumetric
analysis

Stability and
flow test

Graphical
presentation
of test data

Design
bitumen
content

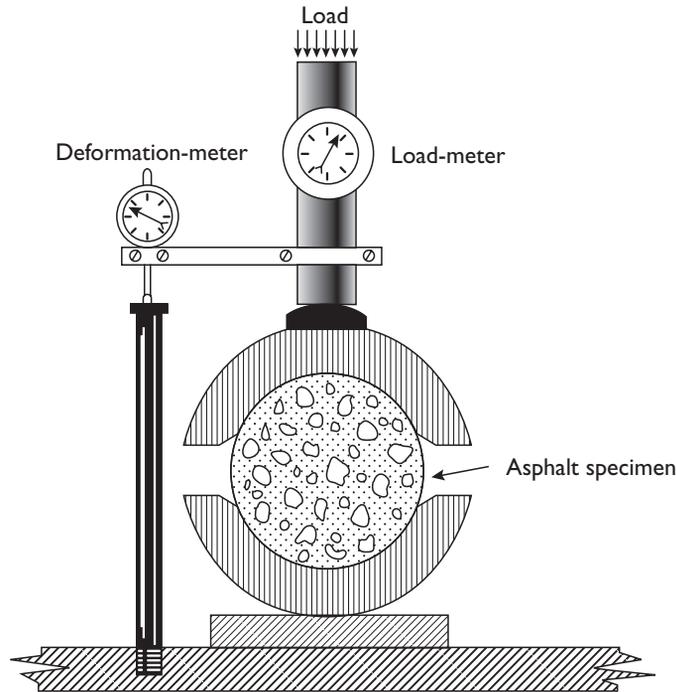


Figure 14.5 Marshall testing machine.

structure is becoming overfilled with bitumen and, as a result, the mix becomes more susceptible to plastic deformation.

Adjustment of
the mix

If one or more design criteria are not met, the grading and/or the quality of the aggregate must be adjusted and new marshall tests carried out until satisfactory results are achieved.

Severe sites

Severe conditions can result from a combination of high temperature, heavy axle loads and slow vehicles. Under slow-moving heavy vehicles, the longer loading time on the pavement results in reduction of the stiffness of the asphalt concrete. This may lead to secondary compaction and plastic failure of the asphalt in the wheel paths. The problem occurs typically on climbing lanes and on approaches to intersections. For asphalt concrete on severe sites, it is recommended that the mix, as determined from the marshall test, be adjusted to secure minimum air voids of 3 per cent at 'refusal density'. Refusal density is obtained by compacting marshall specimens until no further increase in density occurs. Compaction to refusal could be achieved by increasing the number of blows with the marshall hammer to several hundred, but this is not practical. Instead, it is recommended to use a vibrating hammer, which is quicker and more representative of field compaction.

14.4.5 Asphalt mixing

Types of
mixing plants

Although asphalt mix can be produced in situ by hand-mixing, high-quality premixed asphalt can only be produced at stationary or portable mixing plants. Most asphalt mixing plants can be categorized as either a batch facility or a drum-mix facility.

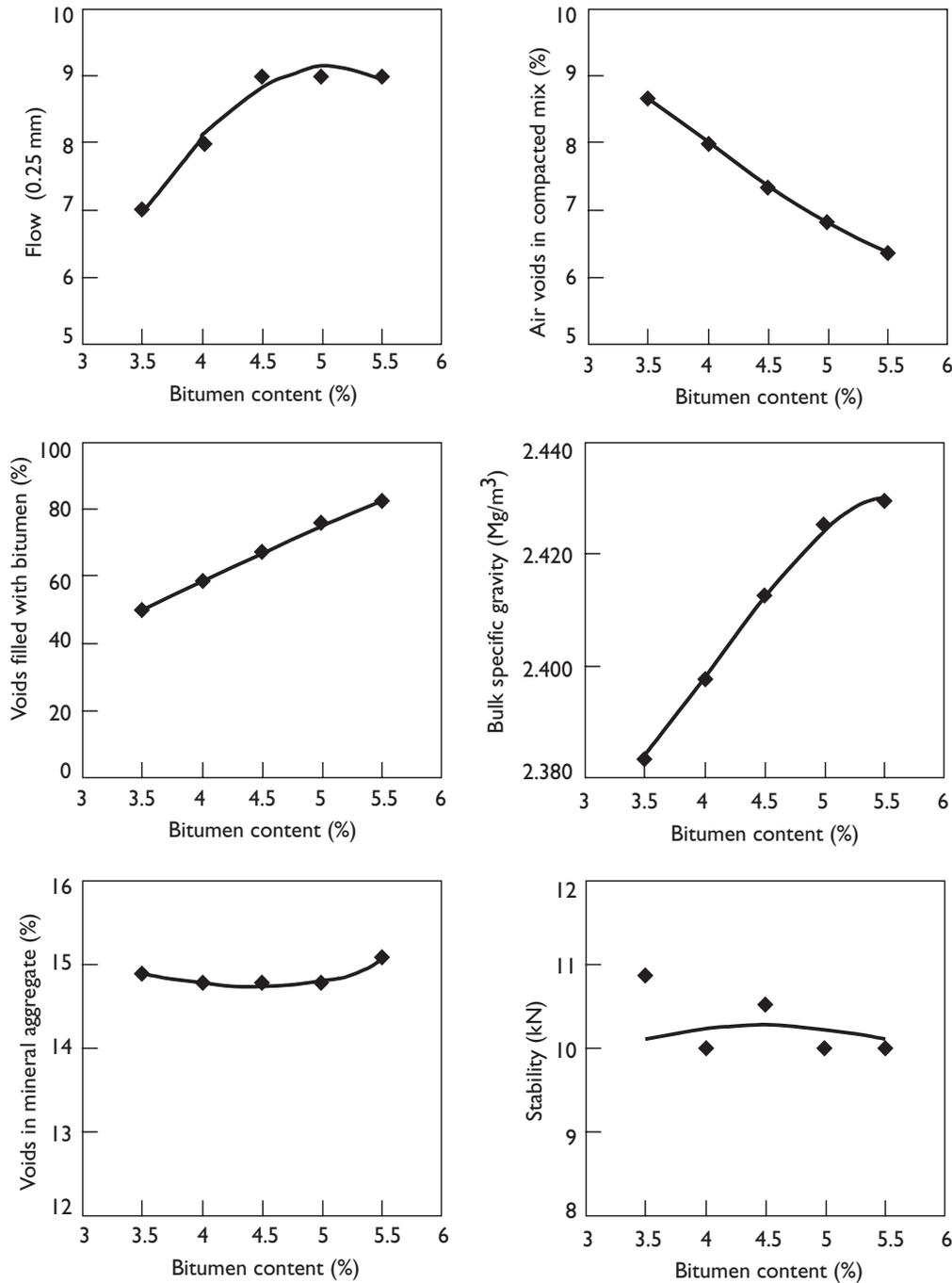


Figure 14.6 Results from a marshall test.

Source: Overseas Road Note 19 (TRL and DFID 2002).

In a batch facility (Figure 14.7), the approximate proportion of aggregates needed is drawn from storage in cold bins onto a conveyer belt leading to an elevator. The elevator delivers the combined aggregates into a dryer.

The dryer is an inclined, rotating, steel drum with a gas or oil-heating unit (heating flame) in the lower end. The combined aggregates enter the dryer at the upper end

Batch mixer

Dryer

Table 14.6 Recommended marshall design criteria for asphalt concrete wearing course

Design traffic load (10^6 ESA) ^a	< 0.4	0.4 – 1	1 – 5	> 5
Compaction level (No. of blows)	2 × 35	2 × 50	2 × 75	2 × 75
Voids in aggregate VMA (%)	Min. 11–16 depending on max. particle size			
Air voids in mix VIM (%)	3.5–4.5	3.5–4.5	3.5–4.5	4.5–5.5
Voids in aggregate filled with bitumen VFB (%)	70–80	65–78	65–75	65–73
Minimum stability (kN)	3.3	5.3	8.0	9.0
Flow (mm)	2.0–4.5	2.0–4.0	2.0–3.5	2.0–3.5

Source: *Overseas Road Note 31* (TRL and ODA 1993).

Note

a The concept of ESA (equivalent standard axles) is explained in Chapter 3.

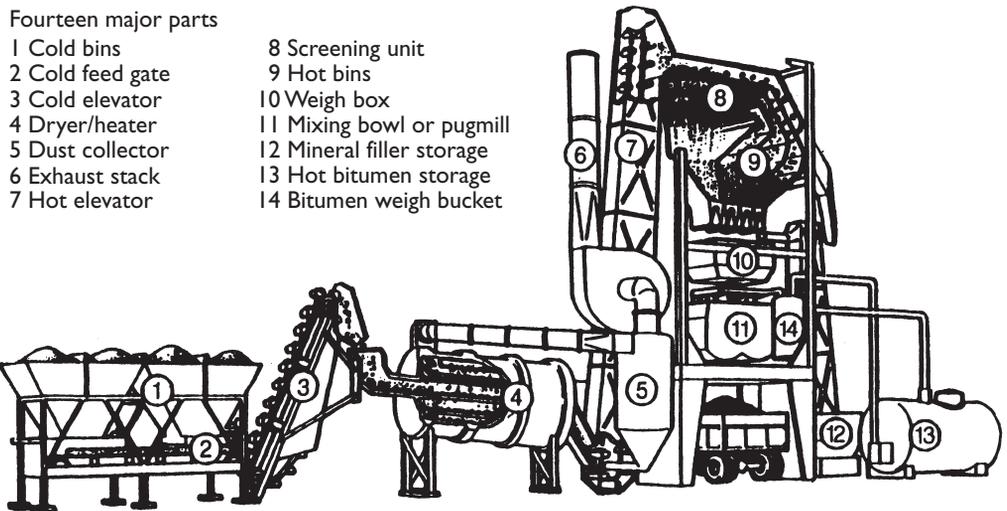


Figure 14.7 Batch mixer.

Source: *The Shell Bitumen Handbook* (Read and Whiteoak 2003).

and are moved against the flame by the rotating action of the drum. As the aggregates pass through the dryer, the moisture is driven off, and the material is heated to a temperature somewhat higher than required for mixing.

Screening unit

The hot, dry aggregates are conveyed up the elevator to a screening unit, where they are separated into several sizes. The different fractions are stored temporarily in 'hot bins' placed below the screens. A prescribed amount of aggregate is successively drawn from the hot bins into a weigh-box located below the bins. Filler at ambient temperature is added from a separate storage bin. At some batch facilities, the heated aggregates are not separated by size into different hot bins. With accurate feeders at the stockpiles, satisfactory grading control can be obtained without this re-screening.

The proportioned aggregates and filler are discharged into a twin-shaft mixer (pugmill) where a measured amount of hot bitumen is added through a spray bar and the materials thoroughly blended. The mixed batch of asphalt is discharged into a truck or transferred to a storage bin.

Mixer

All modern asphalt-mixing facilities are supplied with dust collectors to abate the dust nuisance that otherwise results from exhaust from the dryer. The collected dust is usually returned as filler to the hot aggregate as it emerges from the dryer. Caution must be exercised as these fines may be excessive or not of the right shape or nature, as described earlier.

Dust collector

It is of paramount importance that all ingredients are measured accurately if a high quality asphalt mix is to be produced. Accordingly, all weights and metering devices should be calibrated regularly.

Calibration

The temperature of the asphalt, when it leaves the mixer, should be high enough to get good particle coating, and to ensure proper workability for laying and compacting. However, to avoid burning and to reduce hardening of the binder, close control of the mixing temperature is essential. Typical mixing temperatures are 140–170°C for bitumen penetration grade 60/70, and 150–180°C for penetration grade 40/50. Temperature-measuring instruments should be placed in the discharge from the dryer, in the bitumen storage tank and in the bitumen feeder line. The instruments should be checked frequently.

Temperature

The principal features of a drum-mix facility are that

Drum mixer

- the drying drum also serves as a mixer;
- the output of asphalt mixture is continuous (see Figure 14.8).

A drum-mix facility has no hot aggregate screens or hot bins. The cold aggregate storage bins are supplied with accurate metering devices, which feed a continuous stream of correctly graded aggregates and filler, at a controlled rate, into the drum.

In contrast to a batch facility, the heating flame is placed in the upper end of the drum, where the aggregates are introduced, and the material moves through the drum away from the flame. A bitumen-metering device feeds hot bitumen into the drum at a controlled rate. The binder is introduced midway in the drum to avoid contact with the open flame. From the drum, the hot asphalt mixture is taken by a conveyor to a surge bin, from which it is loaded into trucks.

Drum

A drum mixer is simpler and cheaper to operate than a batch plant. However, drum mixers, and batch plants without hot screens, require a steady flow of aggregates

Comparison

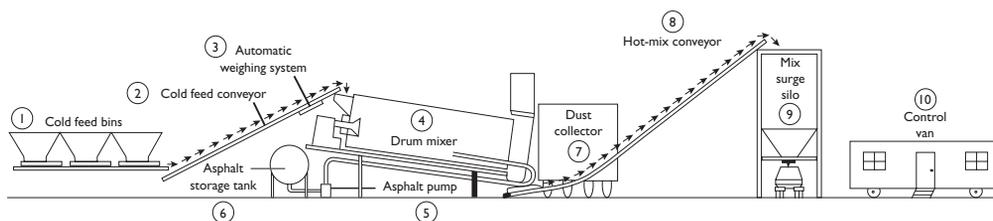


Figure 14.8 Drum mixer.

Source: *The Asphalt Handbook* (Asphalt Institute 1989).

with consistent moisture contents. Furthermore, considerable disruption is caused when a drum mixer is changed from producing one type of asphalt to another.

14.4.6 Asphalt paving

Surface preparation Unpaved surfaces

The riding quality of the pavement surface relies on proper construction and preparation of the foundation.

Unpaved surfaces should be shaped and rolled so that the paving equipment has no difficulty in placing the material in a uniform thickness to a smooth grade. Loose granular particles should be swept from the surface of the road, but care should be taken not to dislodge the aggregate from the surface. The clean surface should be primed with cutback bitumen using a pressure distributor. Under average conditions, the rate of spray should be from 0.9 to 2.3 litres/m². Emulsified bitumen may be used but, under most circumstances, it is more difficult for an emulsion than for a cutback bitumen to penetrate a granular base.

Existing paved surfaces

Existing asphalt pavements should be carefully repaired before receiving a new overlay. Pot-holes should be patched, and the patches should be deep enough to strengthen the base. All bleeding and old unsuitable patches, excess asphalt crack sealer, and loose scales should be removed from the pavement surface. All depressions deeper than 25 mm should be overlaid with a levelling course and compacted. After cleaning with power brooms, a very thin tack coat of emulsified bitumen should be applied to ensure a good bond between the old and new asphalt.

Spreading asphalt

Most premixed asphalt is laid by a paving machine. The asphalt is brought to the paving site in trucks and deposited directly into the paver. The paver spreads the mixture in a uniform layer of the required thickness and shape, ready for compaction.

Asphalt paver

An asphalt paver consists essentially of a tractor and screed unit. The tractor unit provides the motive power through crawlers, or pneumatic tyres, travelling on the sub-base or base. A side view of a paver is shown in Figure 14.9. The mix is dumped into the receiving hopper at the front of the machine. Slat feeders carry the mix back through the control gates to the spreading screws, which distribute the mix in front of the screed unit.

Screed unit

The screed unit is attached to the tractor unit by two long arms that pivot well forward on the paver. When the screed is pulled into the deposited material, it automatically rides up or down seeking the level where the path of its flat bottom surface is parallel to the direction of pull. The thickness of the asphalt mat can be increased or decreased by adjusting the angle between the screed plate and the pull-arm, or by moving the pivot point of the pull-arm vertically. The screed may be controlled manually or automatically. Automatic control systems obtain data from a sensing device riding on a string-line, or from a ski riding on an adjacent lane. The screed unit strikes off the surface, partially compacts the material, and 'irons' the surface of the asphalt mat as it is pulled forward. Most screeds vibrate, but some have a tamper bar that moves up and down.

Compaction

Asphalt is compacted to achieve the optimum density of the mix and provide a smooth riding surface. Compaction is carried out by rollers, which, by their weight or by exertion of dynamic force, compacts the asphalt mat. Self-propelled rollers, such as steel-wheeled tandem rollers, vibrating rollers and pneumatic-tyred rollers are normally used for compaction of high-quality asphalt mixtures.

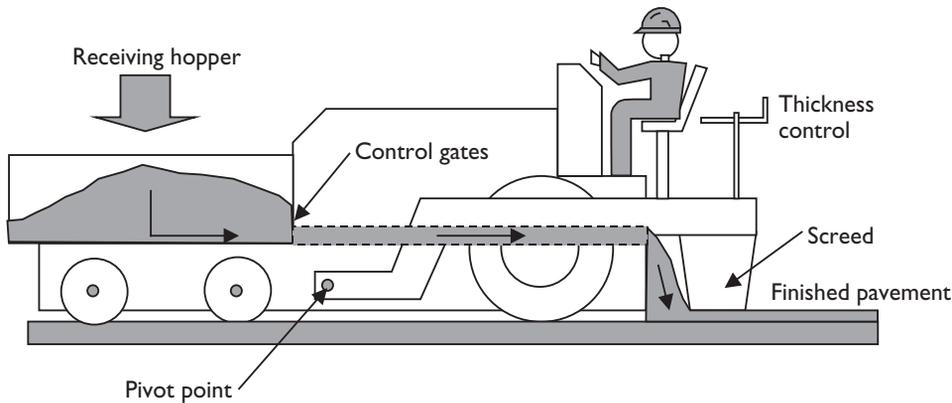


Figure 14.9 Asphalt paver.

Two or more rollers are needed on other than small jobs. Rolling should begin as soon as possible after the hot asphalt has been spread. The rolling operation should start from the lower edge of the lane being paved and progress towards the higher edge. This is because the hot asphalt mixture tends to migrate downhill under the action of a roller. Rollers should move at a slow but uniform speed, with the drive wheel facing the paver. The line of rolling and the direction of rolling should be changed slowly in order to avoid displacement of the mix. Roller wheels should be kept moist with sufficient water to prevent the picking up of material. Heavy equipment, including rollers, should not be permitted to stand on the finished surface before it has cooled.

To obtain good compaction, asphalt concrete should normally be placed in layers with a compacted thickness between twice the maximum aggregate particle size for fine mixes and four times the maximum particle size for coarse mixes.

Control of compaction is achieved by determination of the density of cores drilled from the finished asphalt layer relative to the marshall density. It is recommended that the specified field compaction should be 97 per cent of the marshall density. Other compaction requirements apply to asphalt concrete designed to refusal density.

Number of rollers

Layer thickness

Compaction control

References

- Asphalt Institute (1989) 'The Asphalt Handbook', *Manual Series No. 4*, College Park Md: The Asphalt Institute.
- Millard, R.S. (1993) 'Road building in the tropics', *State of the Art Review No. 9*, London: H.M. Stationery Office.
- Read, J. and Whiteoak, D. (2003) *The Shell Bitumen Handbook*, Manchester: Shell Bitumen UK.
- TRL and DFID (2000) 'A guide to surface dressing in tropical and sub-tropical countries', *Overseas Road Note No. 3*, 2nd edition, Crowthorne: Transport Research Laboratory.
- TRL and DFID (2002) 'A guide to the design of hot mix asphalt in tropical and sub-tropical countries', *Overseas Road Note 19*, Crowthorne: Transport Research Laboratory.

Structural design of asphalt pavements

John Rolt

15.1 Introduction

Aims	<p>The primary purpose of the pavement is to reduce the stresses on the subgrade to such a level that the subgrade does not deform under the action of traffic. At the same time, the pavement layers themselves need to be strong enough to tolerate the stresses and strains to which each layer is exposed. This is the process of structural design. However good the pavement design, the condition of the road will slowly deteriorate with time and traffic. The long-term behaviour of the road will also depend on the maintenance that is undertaken. The aim is to design the road to carry traffic satisfactorily for a specified period of time without needing major structural maintenance (renewal).</p>
Issues	<p>Decisions need to be taken about how much deterioration can be tolerated and on condition that is acceptable at the end of the design period – the terminal condition. Opinions differ on these issues amongst engineers from different countries, and among road users. Whilst engineers are concerned about structural issues, road users will be concerned primarily with the quality of the ride, the slipperiness of the road, congestion and safety.</p>
Structural and functional failures	<p>‘Functional’ failures occur when the road ceases to satisfy the needs of road users; ‘structural’ failures are when the pavement requires renewal because the deterioration cannot be corrected by routine or periodic maintenance. The two are neither the same, nor do they always occur at the same time. Either can occur first. For example, the surface of an asphalt-surfaced road in a dry area can crack quite badly long before the riding quality is affected and before road users begin to complain. Although crack sealing can extend the life, cracking is usually a structural failure that requires an expensive repair. On the other hand, an old road that comprises an unbound base and a simple surface dressing can become very uneven through regular maintenance patching, occurring over many years, although remaining structurally sound.</p>
Life cycle costs	<p>The most common way to balance conflicting aspects of design is to use the principle of minimizing the total cost of the road over its whole life. Total cost means the costs of building and maintaining the road, plus the costs to the users of operating their vehicles and the costs imposed on the rest of society. As the road deteriorates, the road user costs increase. By making realistic assumptions about future maintenance and road behaviour under different maintenance strategies, the road can be designed to minimize total costs over the design period (see e.g. Figure 6.6). Such</p>

an approach requires knowledge of the way roads deteriorate and the effect of this deterioration on the costs of operating vehicles. These issues are discussed in more detail in Chapter 21. When traffic is high, road user costs are also high and, therefore, a higher standard of road is usually provided. When traffic is low, road user costs are also low and it is not economically justifiable to provide a high standard of road.

Methods of pavement design can be subdivided into two main groups:

- methods derived purely from empirical studies of pavement performance;
- methods which make use of the calculated stresses and strains within the pavement (theory), together with studies of the effect of these stresses and strains on the pavement materials (mechanistic behaviour). These are usually called ‘mechanistic methods’, ‘theoretical methods’ or, simply, ‘analytical methods’.

Methods of design

The two methods are complimentary and should always be seen in this way. Empirical methods require some theoretical understanding to help extend them to different conditions, whilst mechanistic methods require empirical information for calibration. Neither method is ideal on its own, but the combination of the two provides a competent basis for design.

15.2 Basic empirical methods

Two empirical methods are described: AASHTO (1986) from the USA and TRL’s *Overseas Road Note 31* (TRL 1993). There are seven key factors that need to be addressed in the structural design process:

Factors

- the terminal condition of the road;
- the strength of the underlying subgrade soil;
- the strength of the pavement layers that are to be used;
- environmental effects;
- traffic;
- maintenance;
- reliability.

The terminal level of deterioration, although important, is not normally a factor that can be chosen by the designer. Usually the level is built into the design method with, for example, a higher level of deterioration being implicitly allowed for lower standards of road. The terminal serviceability is rarely designed for the average level of performance. This is because road performance is highly variable and, therefore, a considerable number of roads would perform very poorly if designs were based on average behaviour. For example, if the average life is 15 years, it is very likely that more than 10 per cent of the roads would last for seven years or less. In most countries this would be unacceptable for all but the very lowest standard of road. Thus, variability is often taken into account implicitly by the authors of design methods by applying different levels of reliability for different classes of road, or explicitly by allowing the road designer to choose a level of reliability, as in the AASHTO method. The designer needs to know the standard deviation of performance in the

Terminal condition and reliability

particular environment for this, but assumptions or informed guesswork normally have to be used.

Strength	Subgrade strength and strength of the pavement layers are important design parameters in all methods. However, most methods allow very little flexibility in the strengths of the pavement layers. Usually these are controlled through rigid specifications. Exceptions to this are sometimes made for specific materials by allowing a trade-off between strength and thickness, but only over a narrow range of strength. For example, weaker materials are sometimes allowable for use in the lower standards of road, but their minimum strength is tightly specified.
Environment and maintenance	Environmental factors are the least understood aspect of design and give rise to the greatest difference between methods. As a result, the transfer of methods between countries can give rise to problems. All methods assume that a certain level of minimum maintenance is carried out.

15.3 Overseas Road Note 31

15.3.1 Principles of the approach

Scope	<i>ORN31</i> is appropriate for roads required to carry up to 30 million cumulative equivalent standard axles (see Chapter 3) in one direction. Although it can also be used for the design of flexible roads in urban areas, considerations such as kerbing, subsoil drainage, skid resistance, etc., are not covered.
Basis of method	<p><i>ORN31</i> is based primarily on</p> <ul style="list-style-type: none"> • the results of full-scale experiments where all factors affecting performance have been accurately measured and their variability quantified; • studies of the performance of as-built existing road networks. <p>Where direct empirical evidence was lacking, designs have been interpolated or extrapolated from empirical studies using road performance models (Parsley and Robinson 1982; Paterson 1987; Rolt <i>et al.</i> 1987) and standard analytical, mechanistic methods (Powell <i>et al.</i> 1984; Brunton <i>et al.</i> 1987; Gerritsen and Koole 1987).</p>
Terminal condition	The terminal condition of the pavement is not defined in <i>ORN31</i> , but the levels of deterioration that are reached by the end of the design period have been restricted to those that give rise to acceptable economic designs under a wide range of conditions. At the end of the design life, the pavement will not be completely worn out, but will be in a condition suitable for renewal by means of an overlay.
Design life	The selected design life for road pavements is commonly between 10 and 20 years.

15.3.2 Traffic

Axle damaging power	In <i>ORN31</i> , the traffic used for design is specified in terms of cumulative equivalent standard axles (ESA), as defined in Chapter 3. <i>ORN31</i> relates the damaging effect of axle loads to a standard axle of 80 kN using the equation
---------------------	---

$$CN_{\text{esa}} = \sum N_P \cdot \left(\frac{P}{80} \right)^n$$

where CN_{esa} is the cumulative number of equivalent standard axles; N_p the number of axles with load P (kN); and n is 4.5.

An exponent value of 4.5 is used in the design method. However, in the AASHTO Road Test (Highway Research Board 1962), from where this relationship was derived, the value of n was also found to depend to some extent on the thickness of the pavement and the terminal condition. There is now considerable evidence that n differs from 4.5 in many situations. This is not considered in *ORN31*, although some methods do take this into account. Note that the *absolute* damage to the pavement caused by the passage of a standard axle depends very strongly on the thickness of the pavement and the strength of the subgrade. It is merely the *ratio* of damage between two axle loads that remains constant, and it is this that underlies the concept of equivalence. An axle load of 80 kN is used as the standard in most design methods, but care is needed since some methods use other values. For example, 100 kN is quite common.

A tandem axle may do slightly more or slightly less damage than two separate axles, depending on the spacing of the two axles in the tandem set, the suspension system of the truck and the structure of the road. Generally, the calculation of ESA can treat tandem axles as two single axles. In some countries, the vehicle population includes a number of heavy vehicles with large single wheels on the main load-bearing axles, rather than dual wheels. These are more damaging to the road than similarly loaded vehicles with dual-wheeled axles. These vehicles need to be considered separately in any analysis.

Tandem axles

The pavement design should be based on the cumulative number of standard axles in the most heavily trafficked lane of the road. For two-lane roads with two-way traffic and no significant difference between the two traffic streams, the design number of standard axles is assumed to be 50 per cent of the total in the two directions. On narrow single-lane roads, traffic channelling can be severe, so the design traffic should be four times the traffic in one direction. The axle-load distribution and damaging effect of heavy vehicles may vary considerably from one country to another. There may also be marked differences between the axle loads of vehicles using different classes of roads. It is therefore strongly recommended that for major roads, unless recent data are available, an axle-load survey of heavy vehicles is undertaken in the area where the road will be constructed and on the class of road in question.

Design traffic

15.3.3 Subgrade strength

The strength of the subgrade is assessed in terms of the ‘california bearing ratio’ (CBR). To estimate the design subgrade strength, it is first necessary to estimate the design moisture content of the subgrade. Moisture conditions under impermeable asphalt surfacings are classified into three categories.

CBR

In Category 1, the water table is sufficiently close to the ground surface to control the subgrade moisture content, and will dominate the subgrade moisture content at the following depths from the road surface:

Category I

- less than 1 m for non-plastic soils;
- 3 m for sandy clays;
- 7 m for heavy clays.

The best and easiest method of evaluating the design moisture content is to take measurements below similar existing pavements during the wet season. Allowance can be made for different soil types because the ratio of subgrade moisture content to the plastic limit is the same for different subgrade soils when the water table and climatic conditions are similar.

Category 2

In Category 2, the water table is deep, but the rainfall is sufficient to produce significant seasonal changes in the moisture conditions under the road. This occurs when rainfall exceeds evapo-transpiration for at least two months of the year. The rainfall in such areas is usually greater than 250 mm per year and is often seasonal. The subgrade moisture condition will depend on the balance between the water entering the subgrade through the shoulders and the moisture leaving the ground by evapo-transpiration when the weather is dry. The design moisture content should be taken as the optimum moisture content given by the standard proctor compaction test (see Chapter 8). Although it is assumed that the surfacing is impermeable, saturated subgrade condition may be anticipated in some cases in both Categories 1 and 2. In such cases, the subgrade strength should be assessed based on saturated CBR samples.

Category 3

In Category 3, the permanent water table is deep, and the climate is arid throughout the year. Such areas have an annual rainfall of less than 250 mm. The moisture content of the subgrade is unlikely to exceed the optimum moisture content given by the standard compaction test, and this moisture content should be used for design purpose.

CBR tests

The compaction properties of the subgrade soil are determined by carrying out standard compaction tests in the laboratory. Samples are compacted in CBR moulds to 100 per cent of the maximum dry density achieved in the compaction tests. The CBR samples are left sealed for 24 h before being tested to remove any pore water pressure induced during compaction. For saturated CBR measurements, the compacted samples are immersed in water for four days before being tested.

Design CBR

ORN31 does not specify the number of samples that should be tested. This depends on the variability of the subgrade and the variation of groundwater level. It is recommended that the design CBR be taken as the value that is exceeded by 90 per cent of the test results. The best method of obtaining this value is to plot a cumulative frequency distribution requiring a minimum of seven results. If the test results change significantly over sections of the road, separate design strengths should be calculated for each section. If equipment for performing CBR tests is not available, a less precise estimate of the design subgrade strength can be obtained from Table 15.1.

Dynamic cone penetrometer tests

In areas where a new road is to be built on the same subgrade as an existing road, and the water table is located at the same depth, direct measurement of the subgrade strength may be made by use of a 'dynamic cone penetrometer' (DCP) (TRL and DFID 1999).

15.3.4 The pavement design catalogue

Design approach

The designs are presented as a catalogue of pavement structures. Each structure in the catalogue is applicable to a small range of traffic loadings and subgrade strengths. The level of reliability under most circumstances is typically around 95 per cent. The traffic increases by a factor of about two from cell to cell. Such

Table 15.1 Estimated minimum design CBR values in ORN31

Depth of water table from formation level ^a (m)	Minimum CBR (%)				
	Non- plastic sand	Sandy clay $PI^b = 10$	Sandy clay $PI = 20$	Silty clay $PI = 30$	Heavy clay $PI > 40$
0.5	8–14	8–14	3–4	3–4	2
1.0	15–29	8–14	5–7	3–4	2
2.0	15–29	15–29	7	5–7	3–4
3.0	>25	15–29	8	5–7	3–4

Source: TRL (1993).

Notes

The table is not applicable for silt, micaceous, organic or weathered clays.

a The highest seasonal level attained by the water table should be taken.

b PI is plasticity index.

a factor is similar to the standard deviation of performance that can be expected from roads that are constructed to high quality. This means that if the engineer chooses a design from the next cell of the catalogue (i.e. where the traffic is higher), it is equivalent to increasing the level of reliability to over 99 per cent. Similarly, if a lower class of road is required, where a reliability level of about 75 per cent is acceptable, the engineer can use the design in the next lower traffic category. However, if construction standards are variable (i.e. with a higher standard deviation), the levels of reliability will be different.

A capping layer is used on subgrades with a design CBR of 2–4 per cent and, in some cases, also on subgrades with a design CBR of 5–7 per cent. For subgrades with CBR less than 2 per cent, special treatment is required that is beyond the scope of the design method.

Figures 15.2 and 15.3 show two design charts from the catalogue. A key to the charts is given in Figure 15.1. Inputs to the design charts are the cumulative number of equivalent standard axles (CN_{esa}) and the subgrade design CBR.

Capping layer

Design charts

15.4 The AASHTO method

The AASHTO design method was developed from the results of the AASHO Road Test conducted during 1959 and 1960 (Highway Research Board 1962), and is probably the most widely used method worldwide. A number of important concepts emerged from the Road Test that are used in other design methods and are summarized here.

A pavement rating system was developed as part of the Road Test. A broad panel of road users was asked to drive across a variety of different roads and indicate their opinion of the conditions on a scale between 0 (poor) and 5 (excellent). The average rating obtained for each road was called the 'present serviceability rating' (PSR). This was then correlated with objective measurements of road roughness, rutting, cracking and patching, of which roughness was the dominant factor. The objective measure is called the 'present serviceability index' (PSI) and is defined as

PSI

$$PSI = 5.0 - b_1 \log R - b_2 RD^2 - b_3 (C + P)^{0.5}$$

Traffic class	$10^6 CN_{esa}$	Subgrade class	CBR (%)
T1	<0.3	S1	2
T2	0.3–0.7	S2	3–4
T3	0.7–1.5	S3	5–7
T4	1.5–3.0	S4	8–14
T5	3.0–6.0	S5	15–29
T6	6.0–10	S6	30+
T7	10–17		
T8	17–30		

Figure 15.1 Key to the design charts in ORN31.

Source: TRL (1993).

where R is the roughness (inches per mile); RD the rut depth (inches); C the cracking (per cent); P the patching (per cent); and b_1, b_2, b_3 are coefficients.

Terminal PSI

In the AASHTO design guide, the terminal PSI is chosen by the designer. A terminal PSI of 2.0 is recommended for rural roads in the United States, but values as low as 1.5 have been used elsewhere.

Traffic

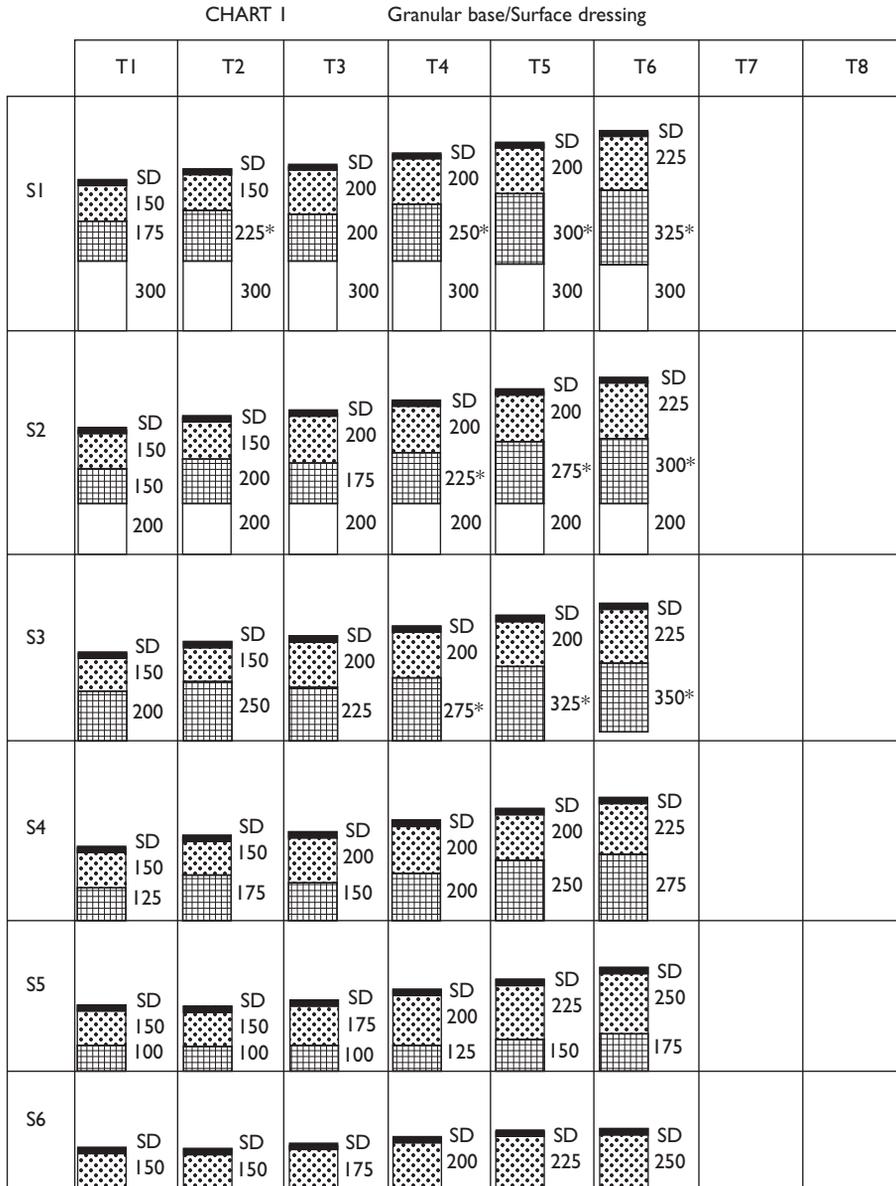
Traffic is considered in terms of equivalent standard axles, as described earlier.

Subgrade strength

Subgrade strength is correlated with resilient modulus. One of these values is estimated for each month (or period of the year), and the values weighted using a nomogram that ensures the functional dependence of road performance and subgrade strength is taken into account properly. A simple average is wrong in principle and should never be used. Some designers choose to use the subgrade strength during the wettest month or when the subgrade is at its weakest to provide a factor of safety, but this is not recommended since reliability and risk are dealt with in the AASHTO method in another way.

Material properties and structural number

Unlike most empirically based design methods, the AASHTO method takes account of variations in material properties and allows the overall thickness to be reduced as the strength of the materials increase above the minimum values required by the specifications. The method employed is based on the structural number



Notes

- 1 * Up to 100 mm of sub-base may be substituted with selected fill provided the sub-base is not reduced to less than the base thickness or 200 mm whichever is the greater. The substitution ratio of sub-base to selected fill is 25 mm : 32 mm.
- 2 A cement or lime-stabilized sub-base may also be used.

Figure 15.2 Design chart from ORN31.

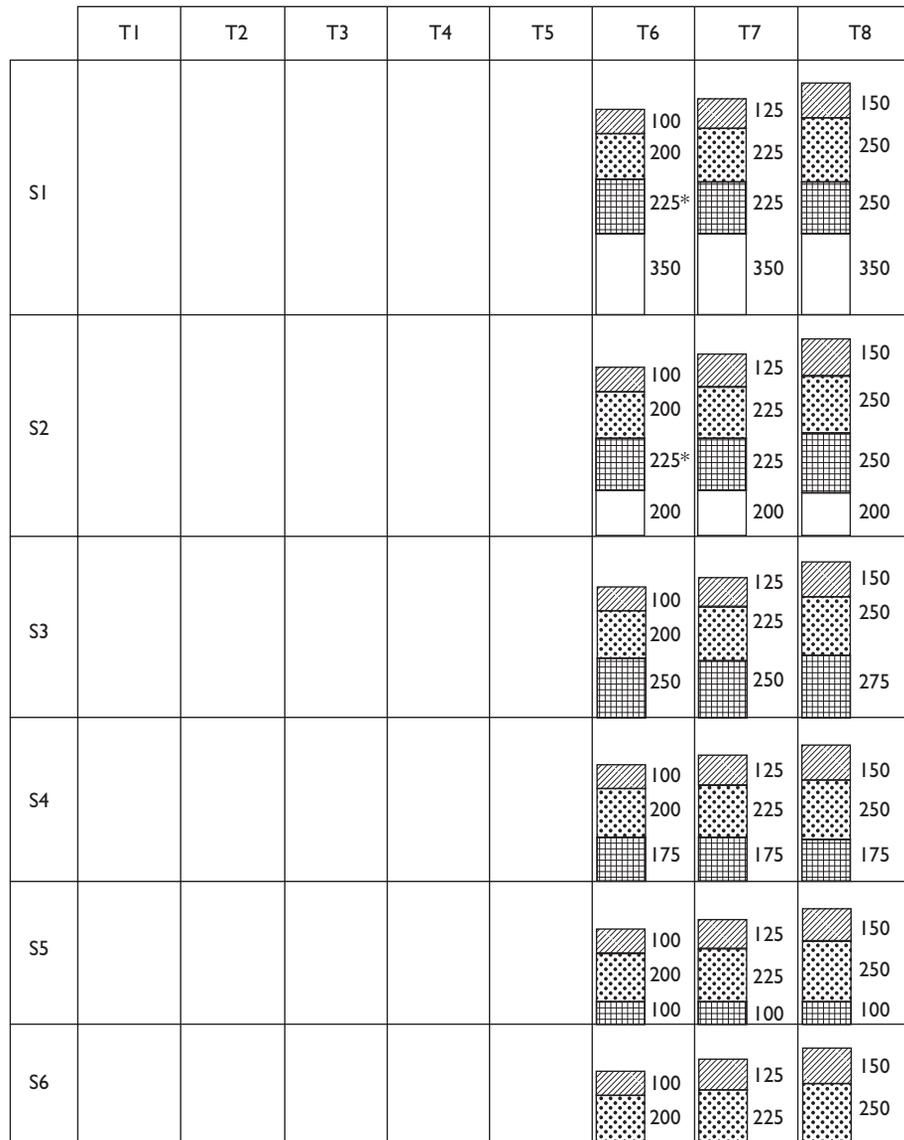
Source: TRL (1993).

principle. The structural number is given by

$$SN = \sum_i a_i \cdot h_i$$

where a_i is the strength coefficients of layer i ; and h_i the thicknesses of layer i (inches).

CHART 5 Granular base/Structural surface



Notes
 1 * Up to 100 mm of sub-base may be substituted with selected fill provided the sub-base is not reduced to less than the base thickness or 200 mm whichever is the greater. The substitution ratio of sub-base to selected fill is 25 mm : 32 mm.
 2 A cement or lime-stabilized sub-base may also be used.

Figure 15.3 Design chart from ORN31.

Source: TRL (1993).

The strength coefficients are related to the standard material strength tests. For example, a relationship has been derived between the strength coefficient for a crushed stone base (a_2) and its CBR value. Similarly the value of a_2 for cement-stabilized materials has been related to unconfined compressive strength. Examples are given in Table 15.2. The coefficient for asphalt materials is very temperature dependant. Lower values should be used in hotter climates.

Table 15.2 Typical strength coefficients

Layer type	Comments	Coefficient
Surface treatment	$h_1 < 30$ mm	$a_1 = 0.1-0.2$
Asphalt mix surfacings	$h_1 < 30$ mm Low stability and cold mixes	$a_1 = 0.20$
Asphalt mix base	$h_1 > 30$ mm	$a_1 = 0.412 \log_{10}(E_1/1000) + 0.246$
	Dense graded with high stiffness	$a_2 = 0.32$ For high temperatures use $a_2 = 0.412 \log_{10}(E_1/1000) + 0.246$
Unbound granular base	CBR > 40	$a_2 = (29.14 \text{ CBR} - 0.1977 \text{ CBR}^2 + 0.00045 \text{ CBR}^3) \times 10^{-4}$
	CBR > 70, on a cemented sub-base	$a_2 = 1.6(29.14 \text{ CBR} - 0.1977 \text{ CBR}^2 + 0.00045 \text{ CBR}^3) \times 10^{-4}$
Stabilized base	Lime or cement	$a_2 = 0.075 + 0.039 \text{ UCS} - 0.00088(\text{UCS})^2$
Sub-bases	Granular	$a_3 = -0.075 + 0.184(\log_{10} \text{ CBR}) - 0.0444(\log_{10} \text{ CBR})^2$
	Cement-stabilized UCS > 0.7 MPa	$a_3 = 0.14$
	Cement-stabilized UCS < 0.7 MPa	$a_3 = 0.11$

Notes

UCS, unconfined compressive strength in MPa at 14 days. E is the resilient modulus by the indirect tensile test in MPa.

Environmental factors are accommodated by weighting the traffic by a constant called the 'regional factor' (R). Thus, in hot and arid areas where the traffic does less damage to a pavement, it is assumed that a design suitable for a lower level of traffic is adequate and, therefore, the regional factor is low. In wet areas, the traffic is assumed to be more damaging, so the regional factor is high. Values range from 0.2 in arid areas to 5.0 in wet areas. Despite the importance of the regional factor in the design process, no detailed guidance is given for selecting its value. Various methods have been used in the United States but, essentially, they are either based on 'engineering judgement', or they are a means of calibrating the AASHTO design method so that it agrees with design charts already in use in the particular region. These approaches are valid provided the behaviour of pavements in the region is known.

The Road Test demonstrated quite dramatically the variability that can be expected in the performance of pavements. For example, the range of traffic carrying capacities of notionally similar pavements designed to carry about 10^6 equivalent standard axles was between 2×10^5 and 5×10^6 . The design method allows a choice of different levels of reliability selected on the basis of the class of road being designed and the policy of the particular road administration.

The AASHTO design equation relates traffic carrying capacity in terms of standard axles (weighted by a regional factor) to structural number, subgrade

Regional factor

Variability, risk and reliability

The design equation

strength expressed in terms of resilient modulus, original PSI, the PSI value selected to define the terminal condition, and the level of reliability desired. At first sight the equation looks formidable, but is actually quite straightforward:

$$\log W_{8.2} = Z \cdot S_0 + 9.36 \log_{10}(SN + 1) + \left[\frac{\log_{10} [(PSI_0 - PSI_f)/(4.2 - 1.5)]}{0.4 + (1094/(SN + 1)^{5.19})} \right] + 2.32 \log_{10} M_R - 8.27$$

where $W_{8.2}$ is the cumulative weighted traffic over the design period; Z the normal deviate (e.g. $Z = -1.65$ for 95 per cent reliability; $Z = -1.04$ for 85 per cent reliability; and $Z = 0$ for 50 per cent reliability); S_0 the standard error of traffic prediction and performance prediction, typically in the range 0.35–0.45; SN , the structural number; PSI_0 , the initial PSI; PSI_f , the final PSI; and M_R , the resilient elastic modulus of subgrade (imperial/US units throughout).

Summary

Thus, to summarize, the AASHTO method requires the following steps:

- 1 Estimate of the total axle loading in equivalent standard axles over the design life.
- 2 Multiply the traffic by the regional factor.
- 3 Estimate the subgrade strength in terms of elastic modulus (note this is 1500·CBR in pounds per square inch, the appropriate imperial unit) weighted as necessary for monthly conditions.
- 4 Select serviceability loss ($PSI_0 - PSI_f$) typically between 2.0 and 3.0.
- 5 Select reliability level and standard deviation of local performance ‘model’ (this is related to construction quality but is usually assumed to be in the range 0.35–0.45 on the $\log W_{8.2}$ scale).
- 6 Use a nomogram to find SN , or use the equation iteratively.

The final step is to choose the various pavement layers and their thicknesses to provide this value of SN . This is done by a simple additive procedure. First, the design procedure is used to calculate the SN of the pavement that would be needed if the base were the subgrade. This will produce the minimum required thickness of asphalt surfacing to protect the base. Next, the design process is used to calculate the SN required to protect the sub-base. Since the contribution of the asphalt surface is known already, the minimum thickness of the base can be calculated. This process is repeated until all layers have been dealt with.

Revision

At the time of going to press, a revision of the AASHTO guide is being drafted that, for the first time since 1960, is expected to incorporate new design and performance equations that are likely to be based on the performance of road sections from the long-term pavement performance programme of the Strategic Highway Research Program (SHRP). This analysis is expected to produce design and performance models based on empirical-mechanistic methods similar in principle to those used in HDM-4.

15.5 Theoretical-mechanistic design

15.5.1 General principles

Methods of pavement design which are based primarily on the results of full-scale experimental road trials are very reliable if used under conditions similar to those under which the experiments were carried out. However, they give little guidance on design if conditions are significantly different. Full-scale trials take a long time and can be expensive. This has helped to drive the development of more theoretical techniques for design based on an understanding of the stresses and strains in pavements and the behaviour of the materials when subjected to these stresses and strains. Such methods can be very powerful, but they have a number of limitations which need to be understood if they are to be used effectively.

Introduction

These theoretical-mechanistic techniques depend on the following general principles:

Principles

- The stresses and strains within the road pavement that cause deterioration can be calculated for a range of different conditions.
- The behaviour of the pavement materials, when exposed to these stresses or strains, can be predicted.
- The critical or allowable values of these stresses or strains, for particular levels of performance, can be identified.

Linear elastic theory can be used to calculate the stresses and strains, with the assumptions that each layer is

Theory

- a linear elastic medium (stress = $K \times$ strain, where K is constant);
- continuous (no discontinuities);
- homogeneous (the properties are the same at all locations);
- isotropic (the properties are the same in all directions).

Some of these assumptions are only approximations. For example, most pavement layers contain discontinuities, such as cracks or voids. They are certainly not very homogenous – density and moisture content will vary from place to place. But most importantly, they are certainly not linear.

The materials' response to stress is usually measured under controlled conditions in the laboratory. Note that this approach is empirical (i.e. experimental), rather than theoretical: there are no fundamental theories about the behaviour of the materials. For example, deformation of a particular material under stress can be measured, as can the dependence on temperature, but it is not possible to calculate how different materials will behave. Thus, the development of mechanistic methods makes use of laboratory rather than field studies. Furthermore, it is also very important that the laboratory tests are carried out under the same conditions that the materials experience in the real road. It is particularly difficult to match field conditions in the laboratory in many cases, and this creates a number of problems, some of which can be quite serious. For example, the behaviour of bituminous materials is very time dependent. The performance of the materials depends on the rate that loads are applied; the basic properties of the materials also change slowly as the material ages.

Material
behaviour

Table 15.3 Typical values of elastic modulus for pavement materials

Material type	<i>E</i> -modulus (MPa)
Asphalt concrete	1,500–8,000
Bituminous macadam	1,500–7,000
Strong cement-stabilized gravel	4,000–12,000
Average cement-stabilized gravel	2,000–7,000
Crushed stone	150–500
Graded natural gravel	100–350
Natural gravel	75–250

Therefore, the accelerated testing in the laboratory does not, and cannot, reflect the real behaviour of the materials in the road.

Field
calibration

It is therefore vital to calibrate laboratory performance against that obtained in the field. The results of calibration are normally presented as an equation relating the stress (or strain) level with the maximum number of repetitions of that stress or strain that can be applied whilst still ensuring satisfactory pavement behaviour.

15.5.2 Linear elastic theory

Parameters

Linear theory requires only two material parameters, namely the elastic modulus (E) and poisson's ratio (ν). Both E and ν are relatively easy to measure but, in practice, ν is fairly similar for each type of material, so generic values can be used. Values are usually in the range 0.3–0.4. Typical values for elastic modulus are given in Table 15.3.

The calculations of the stresses occurring under a wheel load require details of two of the following three related variables:

- load on the wheel;
- contact area of the tyre;
- tyre pressure.

The wheel load is assumed to be uniformly distributed over a circular area. Thus

$$P = \pi \sigma_0 a^2$$

where σ_0 is the contact pressure; a the radius of the contact area; and P the wheel load. The results are usually computed for a standard wheel load (half a standard axle) of 40 kN at a tyre pressure of 517 kN/m². This gives a contact radius of about 150 mm.

Software

Several computer programs are available for calculating the stresses and strains within a multi-layered pavement. These require as inputs for each layer, the values of E , ν and the layer thicknesses. Programs such as ELSYM5 are readily available at very low cost, whilst others, such as BISAR developed by Shell, have been marketed in particularly user-friendly format at modest cost.

15.5.3 Failure criteria

The two most critical points in a pavement structure are the top of the subgrade (vertical stress or strain) and the underside of any stiff, bound layer near to the top

of a pavement (horizontal stress or strain). When an asphalt layer is the primary load-spreading layer, this critical point occurs at the underside of the asphalt layer. The behaviour of pavements with cement-bound layers is more complicated because cracking of cement-stabilized bases does not indicate imminent failure. Discussion of this is outside the scope of this book. The relationship between the stress or strain level and the number of allowable repetitions of the stress or strain is governed by fatigue laws that are discussed later. Reliability or risk in mechanistic design is usually taken into account through the selection of fatigue criteria.

15.5.4 Materials

15.5.4.1 Unbound materials

The elastic modulus of unbound materials depends on the overall stress to which they are exposed, increasing with increasing stress. For granular materials, it can be described by an equation of the following form:

$$E = K_1 \theta^{K_2}$$

where K_1 , K_2 are constants for the particular material obtained from repeated load triaxial tests; and θ the sum of stresses.

For analysis, thick granular layers are normally divided into several thinner layers in such a way that the modulus is constant within each layer, but decreases for deeper layers, that is, as the stress decreases. For cohesive soils, the non-linearity is such that the elastic modulus decreases with stress. Since subgrades often consist of cohesive soils, the effect of this non-linearity can be particularly important. Specific difficulties occur when estimating the moduli of pavement layers from a back-analysis of deflection data obtained from a falling weight deflectometer (FWD) or similar device.

The E -modulus can be measured in the field using an FWD, as shown in Figure 15.4. The FWD simulates the dynamic wheel load of a moving truck. The deflection bowl under the load is measured using a number of geophones at different distances from the centre of the load. If the thickness of each pavement layer is known, it is possible to calculate the elastic modulus of each layer. Description of the detailed methodology is beyond the scope of this book.

It is the effective modulus of the subgrade immediately underneath the main wheel load that is required for design. It is estimated best from direct methods, although a relatively poor correlation with CBR is frequently used, thus

$$E = 10 \cdot \text{CBR}$$

15.5.4.2 Bituminous materials

Bituminous materials exhibit complex engineering properties. They are not elastic, but visco-elastic. Thus, the strains depend not only on the applied stress, but also on the length of time that the stress is applied. Hence the elastic modulus depends on

Non-linearity

Falling weight
deflectometer

Subgrade

Elastic
modulus

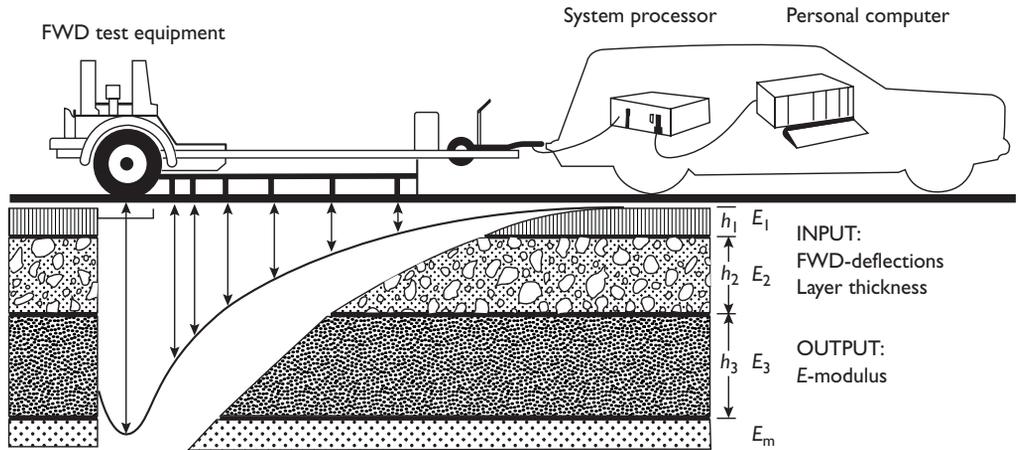


Figure 15.4 Falling weight deflectometer.

the speed of traffic. For example, at 20°C and for fast traffic, the modulus can be as high as 5,000 MN/m² but, on a steep climbing lane where trucks are moving very slowly, it can be as low as 50 MN/m². The elastic modulus is also very dependent on temperature (Figure 15.5). At the high daytime temperatures found in many countries, the modulus can be less than 20 per cent of its value at night. Thus, in the day, the load spreading ability of bituminous materials is severely reduced. The situation is even more extreme in desert areas where, at night, freezing temperatures can occur, giving a modulus 10–20 times higher than during the day. This gives rise to potential failures resulting from thermal as well as traffic stresses.

Fatigue

An asphalt surface does not crack immediately after the road is opened to traffic, but only after many load applications. This is the phenomenon of fatigue. When a specimen is subjected to loading cycles in the laboratory, it can be subjected to either a constant level of stress or a constant level of strain. The results invariably look like those shown in Figure 15.6, where the number of cycles of different stress or strain levels required to produce failure are shown as a function of the stress/strain level. However, traffic consists of a variety of different loads, and these all need to be taken into account in the design.

Predicting fatigue life

Miner's law is used to combine the effect of different stresses or strains. If the fatigue life at stress σ_1 is N_1 and that at stress σ_2 is N_2 , then one application of stress σ_1 uses up $1/N_1$ of life and n_1 applications of load σ_1 uses up a fraction n_1/N_1 . Similarly, one application of stress σ_2 uses up $1/N_2$ of life and n_2 applications of σ_2 uses up a fraction n_2/N_2 . Failure occurs when all of the fatigue life is used up, thus

$$\sum_i n_i/N_i = 1$$

Thus, if the distribution of axle loads (the n s and σ s) in the traffic is known, and the fatigue relationship between σ and the number of cycles to failure (N) is determined

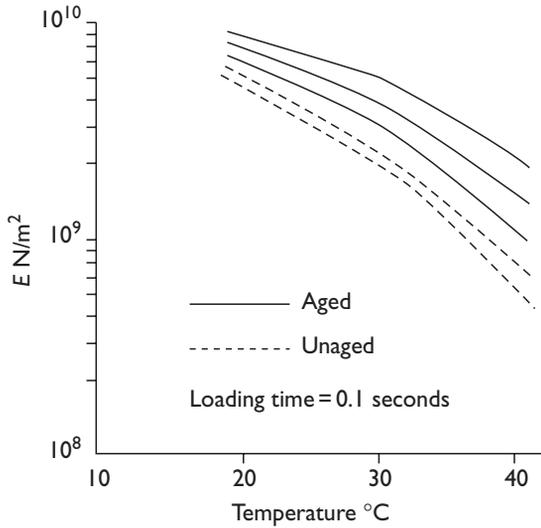


Figure 15.5 Effect of temperature on the elastic modulus of asphalt mixes.

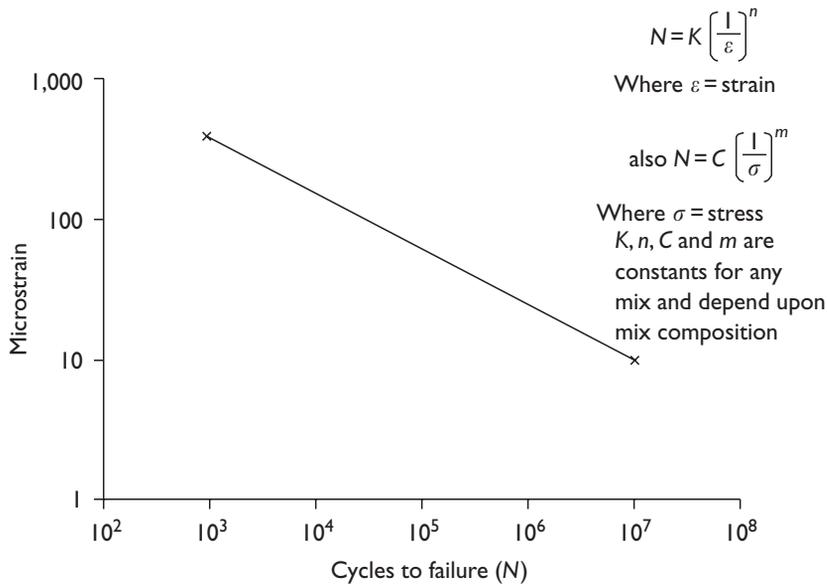


Figure 15.6 Fatigue laws for pavement materials.

from laboratory measurements, the life of the pavement layer can be calculated. However, such calculations predict that all pavements ought to crack within a short time of being opened to traffic. It is clear that the fatigue behaviour in the real road is very different to that in the laboratory. This is because real field conditions cannot be simulated in the laboratory. Bituminous material is visco-elastic and has the ability to 'heal' itself. In the real road, bituminous materials last typically 10–1,000 times longer than would be expected from laboratory studies, so it is necessary to apply

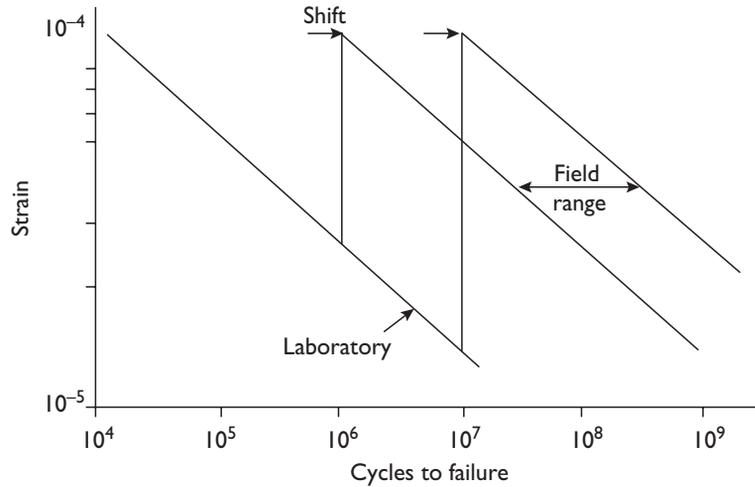


Figure 15.7 Shifting the fatigue line.

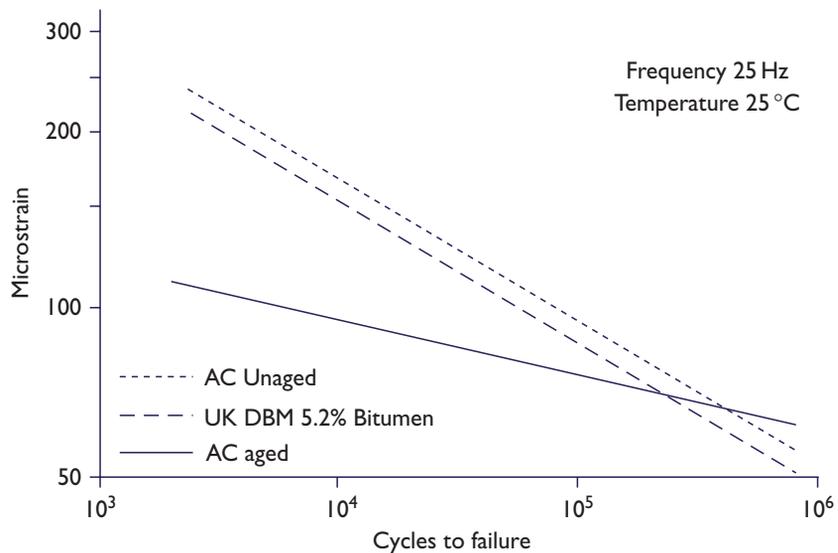


Figure 15.8 The effect of ageing on fatigue.

a ‘shift’ factor to the laboratory data, as shown in Figure 15.7. The situation is more complicated in hot climates because the relatively rapid ageing of bitumen changes the nature of the bituminous materials quite significantly. First, the elastic modulus increases, but the fatigue behaviour also changes (Figure 15.8) as the material becomes more brittle (Smith *et al.* 1990).

Mode of loading

The general fatigue behaviour associated with the constant-strain and constant-stress modes of loading is quite different. When the asphalt layer is thin and the other pavement layers are providing the principal load supporting function, then the mode of loading is constant-strain. When the asphalt is thick and is one of the main load-bearing layers, the mode is closer to constant-stress. The dependence of the fatigue

life on variables such as temperature and mix composition is different in the two modes. Essentially, the fatigue law itself depends on the thickness of the layer.

Empirical data is vital to calibrate laboratory-derived fatigue relationships and, ideally, calibration is needed for different thicknesses of asphalt. The following equation has been developed from research by TRL (Powell *et al.* 1984) for dense bitumen macadam and for 85 per cent reliability:

TRL fatigue relationships

$$N = \left(\frac{5,620}{\varepsilon_{\mu}(r)} \right)^{4.16}$$

where $\varepsilon_{\mu}(r)$ is the horizontal/radial microstrain.

Note that the power factor in this relationship is 4.16. Calibration to local conditions and for other levels of reliability does not involve changing the value of this exponent; it involves a shift in the fatigue line by changing the value of the numerical constant 5,620.

One of the Shell (1978) relationships used by Austroads (1992), the Australian road authority is

Austroads/Shell relationship

$$N = \left[\frac{6,918(0.856V_B + 1.08)}{\varepsilon_{\mu}(r) \cdot E_{\text{mix}}^{0.36}} \right]^{5.0}$$

where V_B is the bitumen content by percentage volume; E_{mix} the elastic modulus of the mix in MPa; and $\varepsilon_{\mu}(r)$ the horizontal/radial microstrain.

This equation is based on a best-fit model and, therefore, represents a fiftieth percentile level of reliability. The advantage of this equation is that different mixes are accommodated through the compositional variable V_B . The effect of temperature and/or loading time on the fatigue law is taken into account largely through the effect of these two variables on the modulus E_{mix} .

The criterion related to the vertical stress on the subgrade is also expressed as a 'fatigue' law. The TRL equation for 85 per cent reliability is

Subgrade criterion

$$N = \left(\frac{15,000}{\varepsilon_{\mu}(z)} \right)^{3.95}$$

where $\varepsilon_{\mu}(z)$ is the vertical microstrain at the top of the subgrade.

The criteria developed in Australia (Austroads 1992) for 50 per cent reliability is

$$N = \left(\frac{8,511}{\varepsilon_{\mu}(z)} \right)^{7.14}$$

Calibration to local conditions or adjustments of the reliability level simply require adjustment of the numerical constant, not the exponent.

Only the subgrade stress level and asphalt fatigue are considered in the design process. This is because the strengths of materials meeting standard specifications, which have been derived from empirical studies, are almost always sufficient to prevent other types of strength-related failure.

Other modes of failure

15.5.5 Complete analysis for design

Once the failure criteria have been defined, the mechanistic method can be used to calculate, for any structure, the stresses and strains at the critical points at any time and season of the year, and for all types of vehicle and wheel configurations. Using miner's law, the accumulated damage created by the entire traffic stream could be computed but, although incorrect in principle, traffic is usually reduced to equivalent standard axles to simplify the calculations. The 'damage' is accumulated until one or other of the fatigue lives at the critical points is exceeded. If this occurs at a level of cumulative traffic that differs from the design traffic, then the layer thicknesses are adjusted and the process repeated until a suitable design is achieved.

15.6 Overlay design

Pavement
evaluation and
diagnosis

The most cost-effective way of extending the life of a road is to make the best use of its residual strength by applying an overlay at the appropriate time, normally before the road shows signs of serious distress. Accurate timing is vital for this to be successful. In reality, few roads are considered for overlaying, or other strengthening, until the condition of the pavement has deteriorated beyond the point where this is a straightforward option. A detailed *evaluation* is then normally required to determine the best solution. This needs to investigate the condition of each layer, including the subgrade, and to determine the cause or causes of deterioration. A reliable diagnostic procedure is vital if the correct rehabilitation treatment is to be identified. An example of pavement evaluation and diagnostic procedures is described in TRL's *Overseas Road Note 18* (TRL and DFID 1999).

Pavement
testing

Pavement evaluation relies upon systematic surveys of the road. This will include a visual survey to identify and to measure defects such as cracks, rutting and other forms of surface deterioration. It will also include measurements of deflection, preferably with an FWD, and road roughness, preferably with an instrument that operates at traffic speeds. DCP tests are also invaluable for measuring the strengths of the pavement layers and their thicknesses. It may be necessary to excavate test pits to obtain samples for testing in the laboratory, although the number of destructive tests such as this should be minimized, and the majority of the necessary information should be obtained in other ways. It is important to recognize that the characteristics of a road can change significantly over very short distances. Therefore, to maximize the chances of making a correct diagnosis and a correct design, it is vital that all point-specific data, for example, deflections, DCP tests, rut depths, are measured at exactly the same place. This will aid diagnosis enormously.

Drainage

For all pavement investigations, the condition of the road drainage should be reviewed and, where deficiencies have weakened the existing pavement, the road drainage should be improved.

Partial
reconstruction

When a road has deteriorated beyond the optimum condition for simply adding an overlay, pavement reconstruction is necessary. The design solution will depend on which layers of the pavement are too weak or too variable in quality to be retained

in their current form, but existing materials should always be used whenever possible. In many cases, deterioration in the form of cracking will have begun at or near to the top of the pavement, with associated problems having slowly propagated downwards. Layers that are deep within the pavement may still be in good condition. Those layers in good enough condition can be retained as sub-base layers without any additional processing. A new base and surfacing may then be required. The variability of a deteriorated pavement will be high, so achieving a high level of reliability will require a relatively conservative approach to design. It may prove more economical to strengthen the old base and surfacing layers in situ with a stabilizing agent, such as cement or bitumen, if the reliability analysis deems this necessary. In this case, only a new surfacing will be required.

However, if the main form of deterioration is occurring deep within the pavement, then full-depth reconstruction may be necessary, again making as much use of the existing materials as possible. In some situations, the ‘problem’ can be ‘buried’ deep enough within the pavement, and be of little consequence, in which case additional layers can be constructed on top of the existing pavement.

The importance of understanding fully the mode or modes of deterioration and their causes cannot be over-emphasized. All decisions rely on the skill and interpretation of the engineer. Under no circumstances should these be left to an automatic process, computer-controlled or otherwise.

The process can be viewed simply as that of designing a new road using the materials that are already in place. Any of the design methods can be used, but the flexibility of the structural number approach of the AASHTO method makes this the most straightforward. The structural number approach can also be used with other design methods. The value of the strength coefficients of the existing layers can be deduced from measurements of strength made during the pavement evaluation phase. The effective structural number of the existing road should be reduced by multiplying the apparent value by a reduction coefficient to take account of the reduced fatigue life of the asphalt materials.

More sophisticated overlay design procedures are available when timely strengthening can be carried out. These rely primarily on evaluating the strength of the road by the back-analysis of ‘deflection bowls’. These measurements are made with an FWD, or similar device. The elastic modulus of each layer is derived which, in principle, indicates its condition. However, back-analysis using mechanistic methods is problematic. The design of overlays requires considerable diagnostic skill, and it is strongly recommended for design purposes that the engineer does not rely solely on this method.

Before the FWD became the instrument of choice for measuring deflection bowls, the simpler and much cheaper ‘benkelman beam’ was used to measure the maximum deflection under a slowly moving loaded-lorry wheel (Smith and Jones 1980). Although much less accurate than an FWD, the benkelman beam deflection value is also closely correlated with the overall ‘strength’ of a pavement, and the beam remains a useful tool for assessing road pavements and for designing the thickness of overlay required to extend the useful life of a pavement (*Overseas Road Note 18*, TRL and DFID 1999). Deflection measurements using a benkelman beam are described in Box 15.1.

Full-depth reconstruction

Role of the engineer

The design process

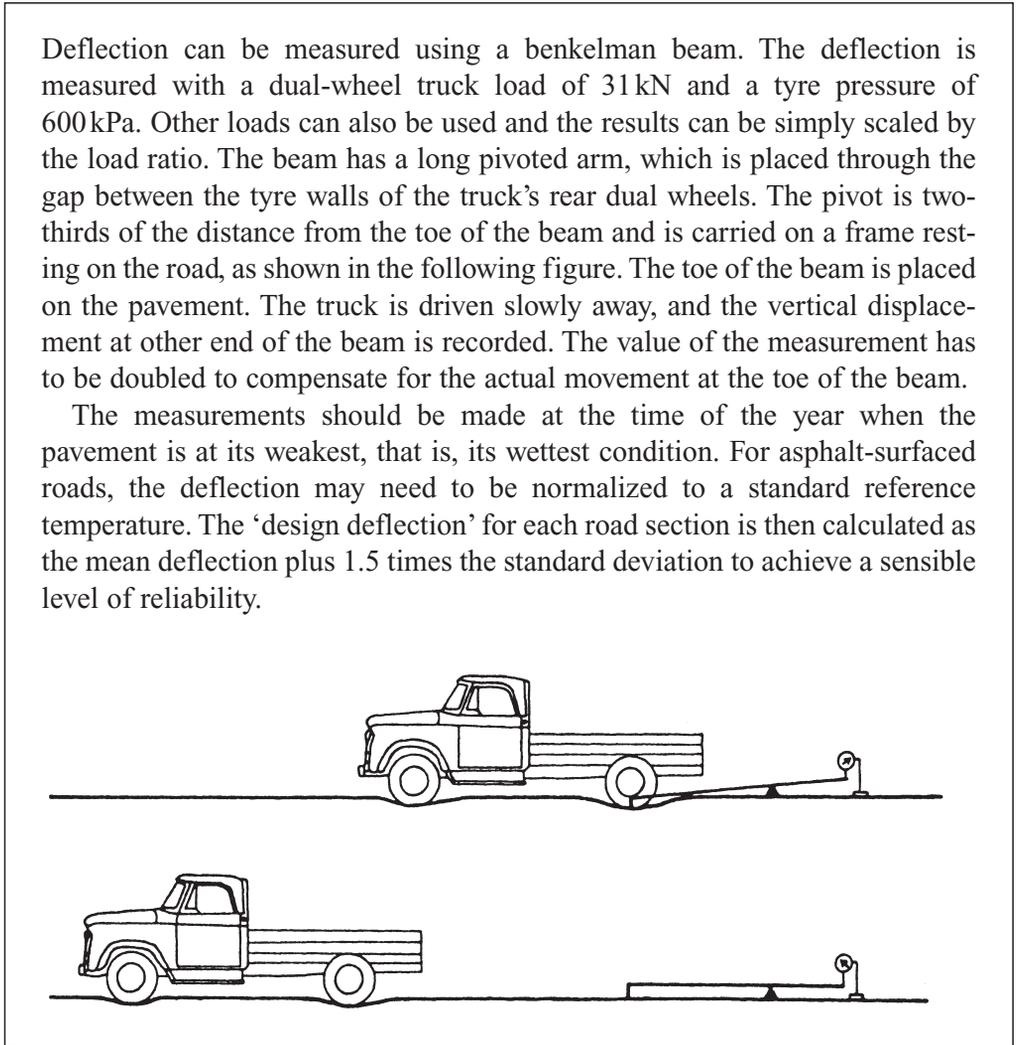
Overlay design

Benkelman beam

Box 15.1 Deflection measurements using a benkelman beam

Deflection can be measured using a benkelman beam. The deflection is measured with a dual-wheel truck load of 31 kN and a tyre pressure of 600 kPa. Other loads can also be used and the results can be simply scaled by the load ratio. The beam has a long pivoted arm, which is placed through the gap between the tyre walls of the truck's rear dual wheels. The pivot is two-thirds of the distance from the toe of the beam and is carried on a frame resting on the road, as shown in the following figure. The toe of the beam is placed on the pavement. The truck is driven slowly away, and the vertical displacement at other end of the beam is recorded. The value of the measurement has to be doubled to compensate for the actual movement at the toe of the beam.

The measurements should be made at the time of the year when the pavement is at its weakest, that is, its wettest condition. For asphalt-surfaced roads, the deflection may need to be normalized to a standard reference temperature. The 'design deflection' for each road section is then calculated as the mean deflection plus 1.5 times the standard deviation to achieve a sensible level of reliability.



References

- AASHTO (1986) *Guide for Design of Pavement Structures*, Washington, DC: American Association of State Highway and Transportation Officials.
- Austrroads (1992) *Pavement Design: A Guide to the Structural Design of Road Pavements*, Surrey Hills NSW: Austrroads.
- Brunton, J.M., Brown, S.F. and Pell, P.S. (1987) 'Developments to the Nottingham analytical design method for asphalt pavements', *Proceedings of the 6th International Conference on the Structural Design of Asphalt Pavements*, Ann Arbor: University of Michigan.
- Gerritsen, A.H. and Koole, R.C. (1987) 'Seven years experience with the structural aspects of the Shell Pavement Design Manual'. *Proceedings of the 6th International Conference on the Structural Design of Asphalt Pavements*, Ann Arbor: University of Michigan.
- Highway Research Board (1962) 'The AASHTO Road Test, Report 5: Pavement Research', *Highway Research Board Special Report 61E*, Washington, DC: National Research Council, National Academy of Science.

- Parsley, L.L. and Robinson, R. (1982) 'The TRRL road investment model for developing countries (RTIM2)', *TRRL Laboratory Report 1057*, Crowthorne: Transport and Road Research Laboratory.
- Paterson, W.D.O. (1987) 'Road deterioration and maintenance effects: models for planning and management', *The Highway Design and Maintenance Standards Series*, Washington, DC: The World Bank.
- Powell, W.D., Potter, J.F., Mayhew, H.C. and Nunn, M.E. (1984) 'The structural design of bituminous roads', *TRRL Laboratory Report 1132*, Crowthorne: Transport and Road Research Laboratory.
- Rolt, J., Williams, S.G., Jones, C.R. and Smith, H.R. (1987) 'Performance of a full scale pavement design experiment in Jamaica', *Transportation Research Record 1117*, Washington, DC: Transportation Research Board.
- Shell (1978) *The Shell Pavement Design Manual*, London: Shell International Petroleum Company.
- Smith, H.R. and Jones, C.R. (1980) 'Measurement of pavement deflections in tropical and sub-tropical climates', *TRRL Laboratory Report 935*, Crowthorne: Transport Research Laboratory.
- Smith, H.R., Rolt, J. and Wambura, J.H.G. (1990) 'The durability of bituminous overlays and wearing courses in tropical environments', *Proceedings of the 3rd International Conference on the Bearing Capacity of Roads and Airfields*, Trondheim: The Norwegian Institute of Technology.
- TRL (1993) 'A guide to the structural design of bitumen-surfaced roads in tropical and sub-tropical countries', *Overseas Road Note 31*, 4th edn, Crowthorne: Transport Research Laboratory.
- TRL and DFID (1999) 'A guide to pavement evaluation and maintenance of bitumen-surfaced roads in tropical and sub-tropical countries', *Overseas Road Note 18*, Crowthorne: Transport Research Laboratory.

Construction



Multiple culvert in India. (Photo: Bent Thagesen)

Contracts and works procurement

Torben Mikkelsen and Tim Waage

16.1 Project execution methods

16.1.1 Options available

Road works can be organized in different ways. These include

Work types

- new construction, including widening or realignment;
- renewals – overlays, pavement reconstruction and rehabilitation;
- maintenance works – routine and periodic;
- operation – road network management.

A number of different procurement options exist that can be used for these. The most common options can be grouped as follows:

- In-house
- Agency
- Contract
- Concession.

16.1.2 In-house procurement

In-house implementation (also known as ‘force account’ or ‘direct labour’) is a well-known method, where all aspects of planning and execution of an infrastructure project are handled exclusively by one party, normally a branch office of a government department. Previously, this approach was common for all types of road works. Today, it is still often used for road maintenance, but its use for new construction and renewal works is declining.

Force account

As the design, construction and supervision are handled by the owner’s organization, disagreement with other parties does not occur, and the personnel involved are familiar with the requirements, policies and procedures related to the project. It is often possible to save time by adopting an in-house approach because works can commence as soon as funding becomes available, even if this is before the design is completed.

Advantages

A disadvantage is that internal cost control is often inadequate, and this tends to result in higher overall project costs than for other project execution methods. There

Disadvantages

is also a tendency towards lax quality control (why check oneself?). It can be difficult to hold anyone accountable for delays and poor quality works, given that the same organization is both executing and supervising the works. There is often a reluctance to replace poor quality works because of the additional costs involved. Work is often implemented in a manner that makes use of available resources (work force, equipment and materials), rather than applying rational and efficient work methods. This is often considered an advantage by the organization, as the project can be used as a ‘buffer’ for the work force, who would otherwise be redundant outside of seasonal or intermittent work activities. However, this does not lead to efficiency.

16.1.3 Agency

Agency agreements

An agency is an organization with delegated authority from the road owner for operation and management of part of the network. Traditionally, agencies have been local authorities that manage part of the road network on behalf of the national road administration. Agencies have, in the past, operated under a framework agreement that sets out activities to be performed, but with considerable discretion to undertake work as they consider appropriate. Typically, activities have included designing, supervising and managing maintenance, improvement works and winter maintenance, and ensuring that the network is maintained to national standards. More recently, agency agreements have moved towards performance-based agreements, rather than simply listing activities to be undertaken.

AGETIP

This is a specific kind of agency arrangement used extensively in francophone Africa. The ‘AGETIP’ is a contract executing agency set up to execute donor-financed infrastructure projects. Some details are given in Box 16.1.

16.1.4 Contracts

16.1.4.1 Works and supervision contracts

Procurement of works contracts

For most project execution methods, with the exception of in-house execution, projects will be constructed under a formal agreement known as a ‘contract’. Under a contract system, the employer enters into an agreement with a contractor, who is often chosen through competitive tendering from a number of bidders. When the process is open to competition from any company, irrespective of their country of origin, this procurement process is often referred to as ‘international competitive bidding’ (ICB). When competition is restricted to local firms, then the process is known as ‘local competitive bidding’ (LCB).

Supervision contracts

To assist in project administration and to supervise the work of the contractor, it is common for the owner to appoint a firm of consulting engineers – in some cases the same firm that designed the project. One advantage of the construction contract is that responsibilities between the three parties (the employer, the contractor and the engineer) are well defined.

Works contracts

Competitive procurement of road works contracts stimulates efficiency and obtains the lowest possible price for a project. The detailed tender documents for these contracts need to include all details about the design, so the method enables

Box 16.1 AGETIP agencies

AGETIP agencies generally have a board of administration composed of well-known figures, but not including government representatives. A general manager is appointed by the board, and line managers and other staff are hired under private sector terms and conditions of service. The agency is set up as a private, non-profit association and pays no taxes. It works on behalf of local authorities who delegate certain functions to the agency. The local government usually reserves the right to select the projects, and the agency then

- recruits consultants to carry out detailed engineering;
- invites bids and awards contracts for supervision and works;
- manages the contracts;
- pays the contractors directly from a special account opened in its own name.

The advantages of the AGETIP are that it

- gets around cumbersome government procurement regulations;
- pays high salaries, and therefore attracts well-motivated, high quality staff.

The disadvantages are that

- the arrangement is not subject to competitive bidding, and can be relatively expensive compared with competitive approaches;
- it is almost entirely dependent on donor funding;
- it probably hampers development of the local consulting industry by creaming off staff and monopolizing all contract execution work for itself under a tax-free operating environment.

Adapted from: Lantran (1990–93).

fair competition between potential contractors. The owner benefits by knowing the financial obligation before commencement of works. The drawback of works contracts is that procurement tends to be more time consuming than other project implementation methods because of the need for the detailed design to be substantially complete. On the other hand, time extensions, claims and the like are safeguarded as a result of the meticulous preparation.

16.1.4.2 EPC contracts

In an ‘engineer, procure and construct’ (EPC) contract, the employer enters into a contract with a single company or entity for both design and construction of the complete project. This is often done on the basis of a fixed price. This type of contract is also known as ‘design and build’ or ‘turnkey’. The advantages of EPC contracts are that the employer has to deal with only one party, who is responsible for all aspects of the works. Total project costs will be known before a final decision

is made to go ahead with the scheme. It is also possible for the works to be started before design has been finalized, which may advance project completion and bring earlier benefits to the employer. Disadvantages are that it can often be complicated to define the terms or content of these projects. This means that it may be difficult to select a company on a competitive basis, as the various companies may offer widely differing terms and conditions. Errors are not easily detected and the employer has less influence on project matters than for other project execution methods. The EPC solution is best suited for projects that are relatively straightforward and well defined. The approach is used widely for projects where specialized expertise is held by a few organizations; for example, standardized housing and industrial complexes.

16.1.5 Concessions

Concession companies

A road concession is the award of a right or a licence to build, own and operate a road for a given period of time. Concessions are awarded to private sector companies to develop and implement a road project, and then operate it long enough to pay back the investment. The road is then normally transferred back to the host government. Concession companies may include private-sector financiers, international contractors, suppliers and other interested parties together with local partners, including government agencies. Organizations such as the European Bank for Reconstruction and Development have been set up specifically to provide financial support to this type of company, either in the form of equity or loan finance.

Types of concession

Different terminology is used for concessions. The term *BOT* stands for ‘build, own and transfer’ or ‘build, operate and transfer’. Other terms in common use include *BOOT* ‘build, own, operate and transfer’, *BOO* ‘build, own and operate’, *BOOST* ‘build, own, operate, subsidize and transfer’, *BLT* ‘build, lease and transfer’ and *DBFO* ‘design, build, finance and operate’. All are types of concessions, and all require the provision of private capital to finance the works.

Advantages

Concessions offer an opportunity for developing and emerging countries to expand their infrastructure without having to finance the construction out of the national budget or through development assistance. Costs are repaid through fees paid by users of the road, normally through tolls. Other forms of ‘shadow’ financing are now also used. Advantages include the following:

- Projects tend to be implemented more quickly than through the public sector.
- Costs are borne by users rather than the tax-payer; new sources of finance can be accessed.
- Risk is passed to the private sector, who are better able to manage it than governments.
- There is an incentive for efficient and innovative construction.

Disadvantages

Despite the opportunities that concessions offer to developing or emerging nations, there is a need to put in place laws and rules that regulate the toll-fees, establish procedures for toll-fee adjustments, and determine the required service level for the road. The legal costs of preparing concession contracts are often high – typically in excess of one million dollars. There are a number of other drawbacks. Investors

are often unwilling to finance road projects, since returns on investment can be marginal and risks high because return periods are long. There is normally a requirement for specific guarantees to be given by governments, often relating to guaranteed minimum toll revenues. Thus, concessions tend to be complex, time and effort consuming, costly to develop and risky.

16.1.6 External financing

The execution of infrastructure projects in developing and emerging countries is often undertaken with some degree of external financial assistance. This assistance may be used for the design and/or the supervision of the works (consultants), the construction of the works (contractors) or a combination thereof. The involvement of external finance is often requested by the employer (owner), being a branch office of a national road authority or municipality. However, international aid and lending agencies often require the involvement of consultants and contractors appointed under ICB, and the application of strict procedures to ensure that the procurement is effective, and meets international standards in terms of efficiency, equal opportunity for bidders and fairness and transparency of the procurement process.

16.2 Types of contract

16.2.1 Forms of contract

The focus of the remainder of this chapter is on works procurement through civil engineering contracts. These can be categorized depending on the payment mechanisms included:

- Unit rate
- Lump sum
- Cost-plus
- Target price.

Payment
categories

The main features of contracts based on these payment types are described below.

In unit rate contracts, the works are broken down under a number of quantified items and listed in a bill of quantities. Prior to works execution, tenderers are requested to quote a unit rate for each of these items. The owner will normally accept the lowest tender, and the unit rate becomes the basis for payment by measuring the work done under each item. The final cost of a unit rate contract often ends up being considerably higher than the tender.

Unit rate

In lump sum contracts, the contractor will be paid a fixed price for completing all works. A difficulty with a lump sum contract is that the contractor's responsibility is assumed to include all work considered incidental to the completion of the contract, whether or not such items of work are included in the contract documents. An experienced contractor should be able to foresee the need for incidental items and allow for their inclusion in the tender. However, problems can occur when the lump sum is deemed to cover risks such as unforeseeable subsoil conditions and even errors in the contract documents, which may substantially change the scope of the

Lump sum

work to be carried out by the contractor. Needless to say, contractors are generally reluctant to carry out works under this type of contract due to the high degree of risk involved. High tender prices can result. Lump sum contracts often result in long disputes on how to interpret the various stipulations in the contract documents, other than for very simple and straightforward works.

Cost-plus

In cost-plus types of contract, the contractor is paid the direct costs of personnel, machines and materials based on salary slips, equipment records, invoices and other related documentation. The only competitive element is for tenderers to quote a percentage mark-up on the direct costs, or a fixed sum, to cover overheads and profit. Such an arrangement is obviously favoured by contractors, as it involves virtually no risks and guarantees a fixed income. Owners, on the other hand, should be wary of cost-plus contracts, as they offer no incentives to contractors to carry out construction in a rational and efficient manner. In fact, the longer the work can be stretched out, the greater will be the payments made. Nowadays, cost-plus contracts are only used for works that are complex, and where quantities are difficult to assess in advance, such as bridge repairs. Alternatively, such works may be undertaken as an EPC contract, provided the requirements for the final works can be clearly defined.

Target price

The target price contract has been developed specifically to add incentives of economy to the cost-plus arrangements. A preliminary cost is estimated and, on completion, the difference between this target and the actual cost is taken into account by calculating a positive or negative adjustment to the mark-up (or fixed sum) according to a pre-agreed formula. The target price approach is an improvement over cost-plus contracts, but often suffers from lack of clarity about how awards and penalties are related to the contractor's performance. Contractors frequently complain that the target is inappropriate and that unforeseeable risks are unfairly allocated causing losses to be incurred.

16.2.2 Types of specification

There are two basic types of specifications that can be used in contracts:

- 'Procedural' (or 'method') specification, where the employer defines details of the work to be carried out (sometimes known colloquially as a 'cook-book' specification).
- 'Functional' (or 'end-product') specification, where the employer defines the result to be achieved by the work in terms of a functional or performance requirement.

Procedural specifications

Procedural specifications have been used traditionally for road works. These reflect the high degree of competence of road administrations, and are relatively easy to specify and to measure. However, they have high supervisory requirements and do little to encourage contractor innovation, since there is little permitted flexibility for changing work methods, designs or materials.

Functional specifications

Functional specifications are used increasingly for road maintenance works contracts, where the amount of supervision otherwise required can present a problem for employer organizations. Defining performance standards in functional terms, such as required road surface friction values, means that supervision requirements are

minimized, since it is only necessary to test the end result. Contractors can then determine the most appropriate way to meet the performance requirement that maximizes the use of their own particular skills, equipment and use of materials. This approach also encourages contractor innovation. The main difficulty with functional specifications is the need to describe and define the functional requirements for all activities. This is simplified by splitting works up into small units. Examples are maintenance works for grass cutting, ditching and surface repairs, where it is relatively easy to define the work in terms of the function of the end product. However, this can lead to problems of co-ordination between activities. Use of these specifications also facilitates the use of small and/or specialist sub-contractors. They may also have more incentive to innovate than firms with more diverse interests because of the need to retain market advantage in their specialist area.

16.2.3 Partnering

Fundamental changes are presently under way in the approach to procurement by contract. In the past, the construction industry has been plagued with claims, litigation and cost over-runs. Less adversarial relationships are now being promoted on engineering projects. These involve the concept of ‘partnering’ and other attempts at making employers and engineering professionals work together to ensure that contracts are undertaken as efficiently and cost-effectively as possible. The aim is to develop, within a contractual relationship, an atmosphere of co-operation and a teamwork approach between employers and contractors, thus reducing conflicts.

This approach is embodied in the *New Engineering Contract* (NEC) (Institution of Civil Engineers 1996). This requires, for example, an undertaking that parties to the contract will act ‘in a spirit of mutual trust and co-operation’. The contract is non-adversarial in nature, and has the aim of promoting good management and quality information for employer and contractor alike. Users of the NEC have found that there are significant advantages and cost savings through the increased openness engendered by the contract. The NEC contains the following options:

- Priced contract with activity schedule or bill of quantities
- Target contract with activity schedule or bill of quantities
- Cost-reimbursable contract
- Management contract.

16.2.4 FIDIC contracts

Contracts are generally complex legal documents which, if disputes arise, are liable to be challenged in arbitration courts. Most countries and donor organizations have therefore adopted standard contract documents, which have been tried and tested for many years under a variety of judicial systems. The most widely recognized forms of standard contracts are produced by the Fédération Internationale des Ingénieurs-Conseils (FIDIC). Other contract systems, such as the earlier British ‘ICE’, are in essence similar to FIDIC. In September 1999, FIDIC issued a new set

Contractual
relationships

New
Engineering
Contract

Standard
contracts

of standard contract forms suitable for a variety of project implementation methods, as follows:

- *Conditions of Contract for Construction*, for traditional contracts where the employer prepares the design (FIDIC 1999a).
- *Conditions of Contract for EPC/Turnkey Projects*, for BOT infrastructure and turnkey projects (FIDIC 1999b).
- *Conditions of Contract for Plant and Design-build*, for most types of works designed by the contractor (FIDIC 1999c).
- *Short Form of Contract*, for small and simple contracts (FIDIC 1999d).

Of these standard contract forms, the most commonly used contract form for bilateral and multilateral financed projects is the *Conditions of Contract for Construction*. The 1999 edition superseded the previous FIDIC 'Red Book' of 1987 (FIDIC 1989, 1992).

Increased use
of FIDIC

FIDIC-based contracts have become increasingly common in recent years for implementation of road projects. The remaining sections of this chapter discuss project execution using the FIDIC system.

16.3 The FIDIC contract

16.3.1 Parties to the contract

There are three parties to FIDIC contracts:

- The 'employer' (owner), who arranges the project financing and the design of the works in addition to employing the 'engineer' and the 'contractor'.
- The 'engineer', who supervises the work of the contractor.
- The 'contractor', whose tender has been accepted by the employer for the works.

Employer

The employer appoints a firm or entity to act as the engineer for the project. FIDIC conditions then envisage that the direct involvement of the employer is limited to the following:

- Issuing the order to commence the works.
- Issuing/approving variation orders that have financial implications.
- Claims.
- Approving new rates for existing items of work that arise from variations in quantities beyond the limits defined in the contract.
- Approval of subletting of any part of the works.
- Approval of any extension of contractual time limits, claims and additional payment.
- Land acquisition and payment matters.

Engineer

The engineer's duty is to ensure that the project is executed in accordance with the terms of the contract between the employer and the contractor, and to supervise the construction to ensure that the works are constructed in accordance with the drawings and the specifications. The previous FIDIC *Condition of Contract for Works of*

Civil Engineering Construction (the ‘Red Book’ of 1987) was founded on the assumption that the engineer will play the role of a generally fair and impartial umpire in the interplay between the employer and the contractor. That concept of impartiality was not always fully recognized or understood by employers, who would expect the engineer to be more on ‘their side’. The current 1999 edition has reduced the contractual impartiality of the engineer and aims more towards ‘making a fair determination’ in cases of disagreements between the employer and the contractor. In countries where the FIDIC concept is new, or in countries that have recently changed to a market economy, problems frequently occur as the employer continues to interfere in day-to-day project activities. This may lead to confusion and blurring of responsibilities, and the engineer’s position can become untenable if important matters are agreed directly between the employer and the contractor.

16.3.2 Selection of supervision engineers

For a project financed by an international lending agency, the employer is often a government department. The appointment of the engineer to supervise the contractor’s work normally follows standard procedures. It is common practice first to establish a long list of consulting firms that might be competent to undertake the supervision. These firms will be invited to express interest in undertaking the assignment and submit relevant information about their background, experience and staffing. In some cases, more detailed information may be requested in a pre-qualification statement of the consultant.

The expressions of interest are evaluated to determine the capability of the firms in question. A shortlist is then prepared containing a limited number (typically six to eight) of the consulting engineers best suited for the job. The firms on this shortlist will be invited to submit a proposal for the consultancy services to meet terms of reference (ToR) prepared by the employer. Normally, separate technical and financial proposals are called for.

The technical proposal will typically have the following contents:

- relevant experience of the consultancy firm;
- comments on the ToR;
- project appreciation/understanding of the ToR;
- description of the technical approach to the services (approach and methodology);
- work plan;
- staffing plan and staff experience, including curricula vitae (CV);
- support facilities, for example computers, laboratories, boring and survey equipment, office, transport, etc.;
- quality assurance plan.

The employer will rank the technical proposal received based on predetermined evaluation criteria, as in the following example.

- Firm’s experience in the field of the assignment: 10 points
- Work plan and understanding to the term of reference: 40 points
- Qualification of staff proposed for the assignment: 50 points

Expression of interest

Shortlisting and terms of reference

Technical proposal

Financial proposal

The financial proposal will normally consist of remuneration in the following areas:

- Consultants' fees.
- Reimbursable expenses – typically consisting of travel costs, per diems, accommodation (housing and office) and transport.
- Miscellaneous expenses – normally consisting of communication costs, office running costs and reporting, and any equipment and vehicles needed for the services.

However, the detailed requirements will depend on the employer and, often, on the standard documents required by the financing institution.

Selection procedures

For many years, the selection of consultants followed a 'quality-based selection' (QBS) method. The parties then negotiated the financial terms for the contract, based on the consultant's financial proposal. However, more recently, there has been a move towards 'quality and cost-based selection' (QCBS). This introduces a competitive parameter to the process, requiring consultants to offer a combination of technical expertise and competitive pricing, against the previous practice where consultants were selected on the strength of their technical expertise only. A typical QCBS approach is summarized in Box 16.2. The above selection procedures are commonly used for World Bank projects. Other financial institutions have similar

Box 16.2 Quality and cost-based consultant selection

This procedure evaluates the technical proposal and assigns points for each proposal. A certain benchmark (minimum number of points) is often set to ensure that only proposals passing the minimum technical requirements will be considered for the overall evaluation with the financial proposal.

The tenderers who pass the minimum technical score may be invited for the opening of the financial proposals. This normally starts with the announcement of the technical score, followed by the opening of the financial proposals and the reading out aloud of the prices.

The lowest price is allocated 100 points, while the others are given the proportional number of points, compared to the lowest price. The most competitive tender is found by combining the technical and financial scores, applying the weight of each element (say 70 per cent of the technical score and 30 per cent of the financial score), to arrive at a total combined score.

The firm of consulting engineers ranked highest on the combined technical and financial scores will be called in for negotiations, mainly on the subject of the technical arrangements. The unit costs are seldom discussed, as they formed part of the competitive tender, although some adjustments may be made to the quantities. If agreement cannot be reached, the firm ranked second will be called for negotiations, and so on. The consultancy contract will then be finalized.

procedures, but there are detailed differences. Sometimes the evaluation can include interviews with the team leader and other specialist staff.

16.3.3 Selection of contractor

16.3.3.1 Basic approaches

In most cases, the selection of the contractor follows a two-stage process involving prequalification and then tendering prior to the award of the contract. This process can be fairly time consuming. It can take 9–12 months to select a contractor for a medium to large project.

Two-stage
tenders

Sometimes the prequalification and tendering processes are combined, with the aim of shortening the time and to prevent the opportunity for collusion between prequalified contractors. Two envelopes are submitted simultaneously, one containing the prequalification statement and one with the tender. In addition, the tender may also consist of two envelopes, one for the tender security and one for the tender itself. With this method, first the envelope with the prequalification statement is opened and evaluated. Tenders who are not prequalified will have their tender envelope returned unopened, while a public opening will be held for the prequalified tenderers. This system is becoming increasingly common. However, the system may not be satisfactory from a contractor's point of view. Many contractors will have the expense of tender preparation without having a chance to assess the competition nor knowing if they are prequalified.

Single-stage
tenders

The methods of selection for both processes are similar and are described below in general terms.

16.3.3.2 Prequalification

With ICB, the selection process normally starts with an invitation for prequalification of contractors in the international and local press, construction journals and through embassies. Interested contractors can respond by submitting typically the following information:

Information
requested

- Past experience
- Previous employers (name and address)
- Present labour force
- Present plant and equipment
- Current and known future commitments
- Financial strength.

The number of contractors seeking prequalification can be very high. By a process of evaluating the information submitted, seeking further clarification and making other investigations, a list is compiled of the contractors who have the requisite qualifications for the job. The successful contractors are notified when the tender documents are ready to be collected, and will receive a set of documents against payment of a document fee.

Evaluation

16.3.3.3 *The tender*

Instructions to tenderers	<p>The contractors will receive ‘instructions to tenderers’, together with the tender documents. These will normally have the following contents:</p> <ul style="list-style-type: none"> • Instructions on how to complete the tender. • Address and time for submission of tender. • Procedures for clarification of issues during the tender period. • Procedures on how to submit alternative tenders, if any and if allowed. • Amount of tender security, normally 0.5 per cent of the estimated tender sum, in the form of a bank guarantee, to be submitted with the tender; the tender security is forfeited in the event that the tenderer does not abide by the contract, if awarded. • Declaration of the tenderer’s obligation to visit the site to acquaint themselves with all local conditions that might affect the construction – such as access, sub-soil conditions, weather, nature of work, availability of materials and local labour supply. • Information about supplementary documents available, for example geotechnical and hydrological investigations, calculations, etc., which are not part of the tender documents. • Declaration that the employer is not bound to accept any of the tenders received, and that the cost of tendering is the responsibility of the tenderers.
Tender period and tender meeting	<p>For ICB road projects, the tender period is normally three months. If, during this period, the tenderers are in doubt about any aspect of the project or the tender documents, clarification must be sought in writing from the engineer. Alternatively, the matter should be brought up at the tender meeting normally arranged by the engineer and the employer for all tenderers.</p>
Tender opening	<p>The tenders must be delivered in sealed envelopes to the employer by the date and time stated in the instructions to tenderers. It is common practice that all tenders are opened immediately after the deadline in the presence of those tenderers who wish to attend. For each tender, the total tender sum is read out aloud, together with a statement on whether or not the tender bond is in order.</p>
Tender evaluation	<p>Next, the engineer will scrutinize the tenders received and prepare a tender evaluation report for the employer. This task comprises</p> <ul style="list-style-type: none"> • evaluating the work programme, construction methods and proposed plant and equipment; • checking whether there are any unacceptable reservations or conditions; • evaluating alternative tenders, if any; • checking the amount of subcontracting and qualifications of subcontractors; • evaluating foreign currency requirements; • checking for unbalanced tender; for example, prices for early work items that are too high, or quantities that are too small; • checking for arithmetical mistakes in the priced bill of quantities; • comparing tender sum with engineer’s cost estimate.
Recommendation	<p>Since a considerable period of time may have elapsed since prequalification, it is often advisable, before a final recommendation is made, to review the two or three</p>

best-placed tenderers, especially in terms of their financial strength and work commitments. If no particular problems are anticipated, acceptance of the lowest tender is almost always recommended. The only exception to this is if the lowest tender is much less than other bids and the engineer's estimate. In such a case, great caution should be exercised, as experience has shown that a loss-making contract is bound to be fraught with problems.

After the employer has accepted the engineer's tender evaluation recommendation, and possibly after negotiation with the winning tenderer, a letter of acceptance is issued. This letter advises the contractor that the tender has been accepted and that an invitation will be issued to sign a contract with the employer. At the time of signing, the contractor is required to produce a bank guarantee, normally amounting to 5–10 per cent of the tender sum, as security for performance of obligations under the contract. Following the signing of the contract, the engineer will issue an order to commence work.

Letter of acceptance and performance bond

16.3.3.4 Subcontractors

The employer will enter into a contract only with the main contractor, selected as described above. Subcontractors may be employed by the contractor under separate contracts, which have no influence on the contract between the employer and the contractor. It follows that the contractor is responsible for the conduct and performance of subcontractors. However, subcontractors need always to be approved by the engineer.

Contractual relationships

Subcontractors can be nominated by the employer directly in certain circumstances. This requires special tendering, and these nominated subcontractors will have to enter into separate contracts with the main contractor. In such cases, the main contractor will still be the responsible party as far as the employer is concerned. Nominated subcontractors are used in cases where the employer wants particular parts of the work to be carried out by specialized companies. Overhead road lighting could be an example of this.

Nominated subcontractors

16.3.4 FIDIC contract documents

The documents comprising a FIDIC contract can be divided into three parts:

Components

- Legal
 - signed contract
 - letter of acceptance
 - tender letter
 - performance bond
 - conditions of contract
 - addenda to the tender documents
- Financial
 - priced bill of quantities
- Technical
 - drawings
 - specifications
 - reference information (frequently not part of the contract documents).

These documents form the basis for tendering. The tenderers are required to fill in the bill of quantities and complete the tender letter based on the legal and technical documents. After the employer and the contractor have signed the contract, the tender documents are referred to as the ‘contract documents’. A brief introduction to the various documents is given below.

Contract agreement

The contract agreement, which is signed by the employer and the contractor, simply confirms that the employer has accepted the contractor’s tender and will pay the contractor as stipulated in the contract. The contractor undertakes to commence and carry out the works in accordance with the contract. The documents comprising of the contract are also identified.

Letter of acceptance

The letter from the employer, informing the contractor that the tender has been accepted, is for the sake of completeness made part of the contract documents.

Tender letter

This letter is submitted by the contractor with the tender and contains the following main items:

- An offer to carry out the works in conformance with the tender documents for a price amounting to the sum of all the items in the priced bill of quantities.
- Acknowledgement of all addenda issued to the tender documents.
- A statement that the tender is valid for a stipulated period of time.
- Acceptance that the tender bond would be forfeited if the contractor should fail to execute the contract when called upon to do so.
- Appendix (with all key data).

Performance bond

A performance bond is a declaration from a bank that it will pay the employer a specified sum of money (normally 5–10 per cent of the tender sum) unconditionally, if the contractor fails to carry out the obligations of the contract.

Conditions of contract

The responsibilities of the parties to a FIDIC contract, as laid down in the conditions of contract, include the following:

- The employer has two principal duties
 - to give possession of the site to the contractor;
 - to pay the contractor what is due within certain time limits laid down in the contract.
- The contractor is under an obligation to execute and maintain the works with ‘due care and diligence’ and subject to the provisions of the contract.
- The engineer will make decisions, issue certificates and give instructions as specified in the contract.

Addenda

Clarifications and modifications to the tender documents, issued by the engineer to all tenderers during the tender period, are included as addenda to the contract documents.

Bill of quantities

The bill of quantities (BoQ) is a list of all items of work, which the contractor is required to carry out. For each item, there is a very short description, and the estimated work quantity is stated. The contractor inserts unit rates and the total estimated price for each individual item in the tender.

Dayworks

The BoQ will normally contain a list of ‘dayworks’. These are used for minor works for which there are no other tender rates, and consist of the contractor’s hourly

rates for various types of labour and equipment. It is normal to quote a mark-up to be paid on top of the purchase price of materials.

Attached to the BoQ are a number of schedules that the contractor fills in as part of the tender. Typically these schedules contain the following information: **Schedules**

- cash-flow projections;
- preconstruction price adjustment indices to allow compensation for fluctuations in the cost of materials and labour;
- the contractor's on-site organization and key personnel;
- proposed subcontractors, if any;
- lists of plant and equipment to be committed to the project;
- land requirements for temporary works; for example, precasting yard and engineer's office;
- local labour requirements, by category.

Detailed drawings are essential for any civil engineering project. For a road project, the drawings normally comprise plans of horizontal and vertical alignment, typical cross-sections, details of services affecting the project, road furniture, layout plans and detailed requirements for structures. **Drawings**

The specifications are supplementary to the drawings and give particulars of the extent of the works, and the quality of materials and workmanship to be attained. This is a very comprehensive document, often running into several hundred pages. **Specifications**

Some reference information is often made available to the tenderers. This can include geotechnical, climatic and hydrological conditions, as well as information on materials sources. The reference document normally contains a disclaimer that the contractor is responsible for the interpretation of the data and is obliged to make independent checks as considered reasonable under the circumstances. The reason for this caution is that the majority of the larger cost claims submitted by contractors stem from unforeseen subsoil conditions. The reference information is therefore excluded from the contract documents in many countries, even though it is made available to the contractor. **Reference information**

Given the large number of documents that form the construction contract, it is virtually impossible to prevent inconsistencies from arising. It is important, therefore, to define in the contract which documents take precedence over others in the case of conflicts. Normally, the legal contract terms take precedence over the financial terms which, in turn, prevail over the technical terms. **Precedence of documents**

References

- FIDIC (1989) *Guide to the Use of FIDIC Conditions of Contract for Works of Civil Engineering Construction*, 4th edition, Lausanne: Fédération Internationale des Ingénieurs-Conseils.
- FIDIC (1992) *Conditions of Contract for Works of Civil Engineering Construction*, 4th edition reprinted with editorial amendments, Lausanne: Fédération Internationale des Ingénieurs-Conseils.
- FIDIC (1999a) *Conditions of Contract for Construction*, Lausanne: Fédération Internationale des Ingénieurs-Conseils.

- FIDIC (1999b) *Conditions of Contract for EPC/Turnkey Projects*, Lausanne: Fédération Internationale des Ingénieurs-Conseils.
- FIDIC (1999c) *Conditions of Contract for Plant and Design-build*, Lausanne: Fédération Internationale des Ingénieurs-Conseils.
- FIDIC (1999d) *Short Form of Contract*, Lausanne: Fédération Internationale des Ingénieurs-Conseils.
- Institution of Civil Engineers (1996) *The Engineering and Construction Contract*, 2nd edition, London: Thomas Telford.
- Lantran, J.M. (1990–93) ‘Contracting out road maintenance activities, Volumes I–IV’, *Sub-Saharan Africa Transport Policy Program, Road Maintenance Initiative*, Washington, DC: The World Bank.

Contract supervision

*Torben Mikkelsen, Bent Thagesen and
Tim Waage*

17.1 Introduction

In a contractual situation, the party wishing to implement a project is usually referred to as the ‘employer’. They will wish to be assured that the project is carried out to the required quality, within the given timeframe and within the budget. The employer needs to make arrangements to protect their interests during construction supervision. Depending on the size and complexity of the project, as well as on the availability of skilled staff, the employer may utilize in-house staff for the supervision, or may engage a consulting firm, or a combination thereof. In developing countries, consultants are often engaged for the key positions and provided with counterparts, who receive on-the-job training to take over ultimately as project managers. In addition, where international financing institutions are involved, consultants are often required (at least) to monitor the use of the loan proceeds.

The employer

The firm or entity appointed by the employer for the purpose of construction supervision is commonly known as the ‘engineer’. The ‘contractor’ is the third party involved in this type of arrangement, and carries out the physical execution of the works. It is the duty of the engineer to ensure that the works are constructed in accordance with the drawings and specifications, and that the contractor fulfils all their obligations under the construction contract. If consultants are involved in the project, they may be appointed as the engineer or the ‘engineer’s representative’, depending on the requirements of the employer and the financing institution.

The engineer
and
contractor

The role of the engineer will depend on the nature of the contract entered into between the employer and the contractor. This chapter discusses the duties and the responsibilities of the engineer under ‘FIDIC’-type construction contracts (FIDIC 1989, 1999a,b), or contracts with principles similar to FIDIC.

FIDIC
contracts

17.2 Supervision organization

17.2.1 Staff and responsibilities

For the purpose of day-to-day supervision of the works, the engineer will appoint an engineer’s representative or, more commonly, a ‘chief resident engineer’ (CRE) or a ‘resident engineer’ (RE). Under the FIDIC conditions, however, it is the engineer that has all the important powers and authorities to act under the terms of the contract. The resident engineer has no authority other than what might be delegated in writing by the engineer.

Resident
engineer

Box 17.1 Typical powers delegated to the resident engineer

The following powers are normally delegated to the RE by the engineer:

- Clarify discrepancies between various contract documents; warn the contractor if progress is too slow; suspend work when immediate action is necessary for protection and security reasons; ensure that project insurances are kept in force.
- Issue further drawings and instructions as required; provide reference points for the contractor's setting out.
- Approve work programmes; order daywork and extra work; look after as-built drawings.
- Preside at site meetings and issue minutes; keep records of all communications with the contractor; deal with third parties and keep records thereof; test and approve the works and keep records.
- Participate and keep records of measurements; forward monthly payment certificates to the engineer; make periodic budget revisions.
- Make recommendations to the engineer on variation orders and time extensions; strive to prevent or at least minimize cost claims; study such claims and make recommendations to the engineer.

Delegation of powers

It is a precondition for an efficient supervision organization that the resident engineer is given authority to act in respect of all routine matters on the site. The engineer is normally not present at the place of work, and it would be impractical to await his decision on urgent site matters. For example, extra expense would be incurred if plant and equipment is idle awaiting a decision from the engineer, and quality problems would arise if the supervision organization were unable to stop faulty works quickly. In international practice, the resident engineer is normally delegated the powers needed to deal with the issues shown in Box 17.1.

Powers not delegated

It is not normal to delegate the more important legal and financial matters to the resident engineer. These typically comprise:

- major variation orders;
- time extension and cost claims; disputes and arbitration;
- certification of contractor's default; completion and maintenance certificates;
- final payment.

RE's staff

The resident engineer needs to be supported by well-qualified staff. Apart from technical competence, such staff should have personal qualities that enable them to work well together with the contractor's personnel to reduce the natural friction between these two parties. The resident engineer's staff normally consists of a combination of:

- Engineers, with responsibility for both technical and administrative matters.
- Supervisors and technicians working on site and in the laboratories.
- Sub-professional staff for administration, accounts, typing, filing, communications and the like.

Recently graduated engineers and technicians are often assigned to supervision duties. Although they need to develop their experience, their technical knowledge alone will prove insufficient against the contractor's seasoned foremen and superintendents on a construction site. It is essential that the resident engineer and key staff are in possession of practical road works experience. By assigning younger engineers to work with more senior staff, they will be given the opportunity to develop their skills. Apart from technical knowledge, senior supervision staff also need to be good administrators, and be acquainted with accounting and contract law.

Expertise

Personal qualities are important, as supervision engineers need to get on with people and deal with the contractor's staff in a fair, reasonable, yet firm manner. This can be difficult in the highly charged atmosphere on a works site. Honesty and professional integrity are essential qualities for resisting offers of monetary awards or other gifts from the contractor.

Personal qualities

17.2.2 Setting up supervision

It is important that the supervision organization is already set up and functional when work is started. Preparations should include a clear organization plan with lines of command and delineation of responsibilities. The number of the staff required will depend on the size and complexity of the project. It is poor economy to have either too small or too large a supervision organization. If the supervision staff is insufficient, the work quality may suffer, and substantial extra costs may be incurred as a result of successful cost claims for disruptions and delays to the contractor's operations. Having too large a supervision staff is of course in nobody's interest.

Requirements

The size and composition of the supervision organization depends, to a large degree, on the project type, and size in terms of the contract value or length of the road. However, often the structure of the organization is dictated by the packaging of the project, where the employer has divided the project into a number of (often equal sized) lots. Each of these may be supervised by a resident engineer under the overall supervision of a chief resident engineer. However, other approaches are also used. Separate contracts may be used for:

Approach

- structures (bridges, culverts, etc.);
- earthworks, sub-base and base;
- pavement;
- road markings and road furniture.

The advantages of this are that specialized contractors often win the contracts and the use of subcontractors is minimized. It may also lead to lower construction costs, as overhead and profit are only added once, rather than separately for the main contractor and subcontractors.

Many issues affect the most appropriate composition of the supervision organization:

Organization

- Complexity of the project – is the project very simple, or does the project call for advanced technologies, which will require considerable skill from both the supervision team and the contractor?

- Logistics – is there a reasonable balance between the travel time and the time spent on site for the supervision team; this applies in particularly to
 - rehabilitation projects where only parts of a long road section will be rehabilitated;
 - new constructions where, say, the bridges are undertaken under separate contracts;
- Staff – if a combination of consultants and counterparts is used, it may be necessary to increase the consultant's organization, depending on the skill level of the counterparts.

An example of a supervision organization for a road project is shown in Figure 17.1.

Facilities

In order to carry out its duties, the supervision organization must be provided with facilities, such as:

- Permanent and mobile offices, fully equipped with telephone, fax, email and internet, radio communication, computers, printers and copying machine.
- Central and field laboratories, fully equipped for all standard testing of soils, pavement materials and concrete.
- Survey equipment, including theodolites, automatic levelling instruments, electronic distance meters, measuring tapes, etc.
- Transport, including four-wheel drive vehicles, as required.

It is common practice for the supervision organization's facilities to be supplied and maintained by the contractor under the works contract. After the project is completed, the facilities may revert to the contractor or may become the property of the employer.

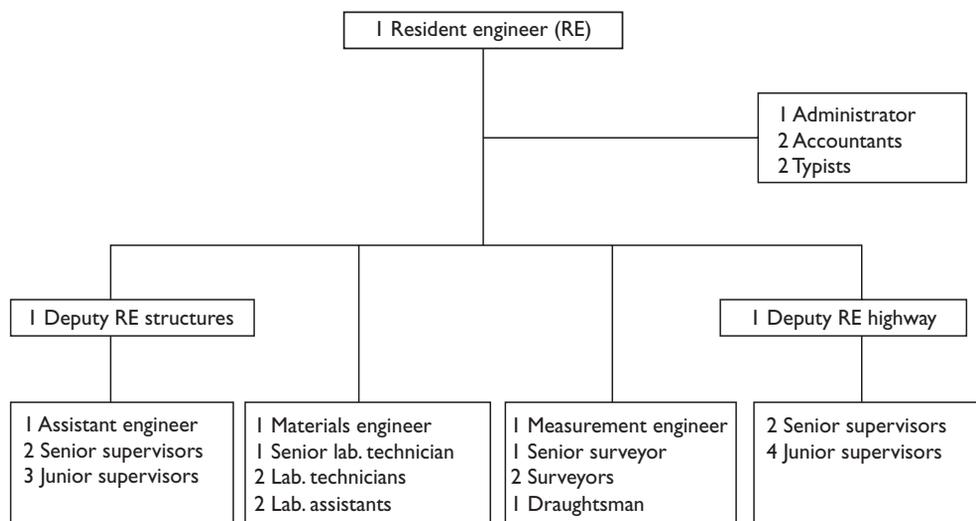


Figure 17.1 Example of a road supervision organization.

17.3 Quality control

Quality control is the principal reason for having a supervision organization on site. To guarantee the quality of works, it is necessary to establish close control over the contractor's workmanship and materials. This task can be divided into the three groups shown in Box 17.2.

Before construction activities can start, testing facilities must be established. This means that permanent and mobile laboratories are ready, laboratory equipment has been acquired and commissioned, test sheets and journals have been prepared, and testing personnel have been employed and received any training necessary. The contractor is often responsible for providing testing facilities for the resident engineer and his staff. It normally takes a long time to establish a functional laboratory. Test equipment may have to be imported, and this often causes delay. Conflicts frequently arise on site, with the contractor wishing to push ahead with earthworks and concreting before the supervision organization is ready to carry out the required quality testing.

The daily operation of the laboratory is normally the responsibility of the consultant's organization. However, in some cases, the contractor is responsible for the operation, and this is monitored by the consultants. This approach is not recommended since it leads to an environment where there can be collusion between the two parties.

There are two types of quality control specifications. One requires testing of the works to check that the end product complies with the specified requirement. It may, for example, be prescribed that the compaction degree, measured at any point in an earth embankment, should be not less than a given value. Where this type of

Activities

Testing facilities

Responsibility

End-product specification

Box 17.2 Quality control activities

Inspection

The works are inspected visually to detect any deviation from the specified requirements. This activity is supplemented by simple in-situ checking; for example, measurement of spacing and cover of reinforcement bars, layer thickness, temperature of asphalt hot mix, concrete slump, etc.

Laboratory testing

Materials as well as the finished product are subject to laboratory testing. For example, checking of asphalt concrete works comprises

- testing of the component materials – crushed rock, sand, filler and bitumen;
- verification of the marshall parameters of the hot mix; and
- compaction control of the asphalt pavement.

Geometric control

Geometric control is required to check compliance with specified dimensions and tolerances, such as length, width, height and camber of pre-cast, pre-stressed girders, and the evenness of finished surfaces such as concrete decks and asphalt pavements.

specification applies, the contractor can choose how to construct the embankment, as long as the specified end-product requirement is attained.

Method specification

An alternative approach places emphasis on the methods used during construction. In this case, for the compaction of an earth embankment, the specification would prescribe a certain number of passes of a particular combination of rollers on a layer of soil of defined thickness and moisture content. The exact method will have to be established based on comprehensive initial trials to ascertain what combination of roller passes, etc., will meet the design criteria. It is important for the author of a specification to decide which type of quality assurance requirements to adopt. A mixture of end-product and method-type specifications can easily lead to confusion during construction, unless it is made clear which requirement takes precedence.

Test methods

In addition to technical requirements, the specifications will normally contain detailed descriptions of test methods, often by reference to internationally acknowledged test designations, such as AASHTO (American Association of State Highway and Transportation Officials), ASTM (American Standards for Testing Materials) and BS (British Standards). Well-defined test methods are mandatory, if disputes are to be avoided during the construction period. For example, it is not sufficient to require that a class of concrete must have a seven-day strength of at least 30 MPa. This requirement is meaningless unless characteristics such as the test specimens' shape (cube or cylinder) and dimensions are defined, along with the rate of load application and the curing conditions.

Extent of control

As works under the FIDIC system must be carried out to the 'satisfaction of the engineer', it is up to the supervision organization to determine the degree of checking to be adopted for any road project. This choice should depend on the type of work to be carried out. Particular care should be taken for operations where the quality of the finished product is heavily dependent on the contractor's workmanship. Strict supervision is also necessary where it is difficult to rectify faulty work, and when work subsequently will be covered up. Pile driving, concrete casting and asphalt paving are examples of work items that require particular attention. In such cases, there should always be an inspector from the resident engineer's staff present. This is referred to as 'standing supervision'. Other work operations, which are less sensitive to workmanship, or can be rectified without undue difficulty, will only require periodic inspection. Earthworks, for example, fall into this category.

Testing frequency

The frequency of testing depends on the quality parameters that require checking. Parameters that are prone to show considerable variation, such as composition of asphalt concrete and compaction of asphalt pavement, should be subject to running control. In such cases, it is common practice to carry out a predetermined number of investigations. Examples would be: one complete extraction and marshall test per 250 tonnes of asphalt concrete produced; or three compaction tests per 2,000 m² of pavement placed. During the start-up period, and whenever quality problems have been identified, the test frequency should be increased. Quality parameters that normally remain fairly constant, such as the los angeles abrasion value of rock from one and the same source, require only occasional checking. The same goes for parameters that can be checked reasonably accurately by eye, for example, the shape of crushed aggregate particles. For materials manufactured under controlled factory conditions, such as steel, cement and bitumen, it is common practice to make only a few random checks and rely otherwise on the manufactures' certificates.

A problem every resident engineer encounters is to decide on the frequency of compaction testing of road embankments. There is no hard and fast rule, as the need for testing will depend on such factors as suitability of the compaction equipment, uniformity of the soil, weather conditions and experience of the contractor's earth-work foreman. As a rule of thumb, it is suggested that each embankment layer be tested at 50-m intervals. Near structures, where differential settlements are likely, the test frequency should be increased considerably to, say, one test per 10-m³ compacted soil.

There will always be pressure to provide test results within the shortest possible time to avoid delay to the contractor. If the results will be available only the next day, the contractor needs to plan operations so that work can be carried out at other nearby locations, while awaiting the results of the testing. For embankment construction, the emergence of nuclear density testing has eliminated the need for time-consuming laboratory testing, and efficiency has greatly improved.

The method adopted for evaluating test results for the purpose of acceptance or rejection of works and materials has far-reaching consequences. For roadworks, quality control is still based largely on absolute requirements and individual spot tests. Typically, the specifications required that an embankment is compacted to at least 95 per cent of the maximum dry density achieved by standard proctor compaction. This means that each and every test result must be equal to or higher than 95 per cent. Whenever a test result falls below this value, the part of the embankment represented by the test should, in principle, be rejected.

The problem with the above approach is that the results of compaction tests for soil are highly variable. This is also the case for many other parameters pertinent to road construction quality. The value of a random test depends on the inherent variability of the material and the errors in measurement, and there can be considerable uncertainties. Most of the test results will be near the true mean, but there will often be some results that are higher and some that are lower. Occasionally, test results will indicate sub-standard material or workmanship, even though the quality is actually good. Conversely, some test results will appear to be in order, even though the work or the material represented by the test is, in reality, substandard.

As a result, the supervision organization would be justified in occasionally accepting a single failed test among many good results. For example, if a particular stretch of embankment is represented by 10 compaction tests, then the whole stretch could be accepted, even if one single test was marginally below the required 95 per cent. However, to avoid creating a bad precedent and accusations of not protecting the interests of the employer, any acceptance of work that has formally failed should, as a rule, be conditional upon further testing showing acceptable results, or instructions to the contractor for remedial action. If, for instance, one compaction test result is somewhat below the specified requirement, the resident engineer may take another test nearby and accept the work if this test is in order. Alternatively, the contractor may be instructed to compact the area in question by, say, another two passes of the roller.

Quality control based on absolute requirements and spot tests does not necessarily ensure a well-defined quality of the product. It is for this reason that a statistical approach to quality control has been adopted within the manufacturing sector, where works and materials are accepted or rejected based on average and standard

Speed of testing

Evaluation of test results

Variability of results

Isolated failed tests

Statistical quality control

deviation considerations. A similar approach is gradually being adopted in the road industry and has, for some years, been common for the quality control of concrete. For other road works, quasi-statistical methods are sometimes used. For example, the specifications for asphalt works may require that the asphalt mixture is compacted so that 'no more than one in any consecutive 10 test results shall be more than two per cent below the specified density'.

Relations with contractor

To maintain good relations, the resident engineer and his staff should make it a rule not to approach the contractor's foremen or subcontractors directly with complaints about the work. If the contractor's project manager is bypassed in this way, they will lose face and authority, which is not in the interest of the supervision organization. The resident engineer should take care not to be too pedantic and rigid with decisions. On the other hand, being fair and reasonable does not mean that works of doubtful quality should be approved. The key supervision issue is that quality considerations must always prevail over the contractor's wish to speed-up work and cut costs.

17.4 Measurement of work

Bill of quantities

The quantities given in the bill of quantities are only estimates. For most items, it is not possible during the project preparation phase to predict exactly the quantities needed for the works. For example, the quantity of unsuitable soil, which has to be hauled to spoil, can only be determined accurately during construction. It is generally considered acceptable if the bill of quantities is accurate to within plus or minus 10 per cent. For payment purposes, it is therefore necessary to measure the works actually carried out.

Methods of measurement

The contract will contain exact descriptions of how the various items of work are to be measured. These descriptions are normally found in the specifications or in a preamble to the bill of quantities. Reference is sometimes made to a standard method of measurement, such as the *Civil Engineering Standard Method of Measurement*, published by the Institution of Civil Engineers in the United Kingdom. Methods of measurement often vary from contract to contract, particularly for items such as excavation of cuttings, ditches and borrow, and haulage to fill. Some quantities, such as sub-base and base, are normally calculated from the theoretical dimensions shown on the drawings, which means that the contractor is not compensated for any over-fill. It is important for tenderers to be aware of these matters, as the method of measurement may influence their pricing.

Notice of measurement

Measurement of completed items of work, according to the standard FIDIC conditions, should be carried out by the engineer, who normally will delegate this responsibility to the resident engineer. Notice must be given to the contractor about the intention to measure work, who must then participate in the measurement and give all necessary assistance to the resident engineer. In the event that the contractor fails to attend, the resident engineer's measurements are taken as correct. It is not unusual, however, for contracts to be written to give the responsibility of measurement to the contractor, in which case it is the duty of the supervision organization to check and approve the contractor's measurement.

Monthly measurements

Monthly measurements are made for the purpose of the contractor's interim payment certificates. Approximate estimates will suffice for work items in progress.

Detailed measurements must be carried out as soon as any item has been completed and before it is covered up.

In addition to the pay items listed in the bill of quantities, the resident engineer should measure and keep records of ‘non-pay items’ for which the contractor may later present claims for extra payment. Such items include

Non-pay items

- large boulders in soil excavation;
- unforeseen settlements in embankment areas;
- slides and cave-ins;
- broken piles;
- public utilities found within the work area.

17.5 Payment to the contractor

It is normal international practice for the contractor to receive an advance payment at the outset of the works to cover mobilization costs. This advance, which is typically 10–15 per cent of the contract sum, is only payable after the contract has been signed and the contractor has provided the required security in the form of a performance bond. A bank guarantee is also required to provide security for the mobilization advance itself. It is important for the supervision organization to make sure that both these securities are valid and formally in order.

Mobilization
advance

During construction, the contractor is entitled to receive regular monthly or quarterly progress payments for works carried out. The procedure is that the contractor submits to the resident engineer a monthly or quarterly statement showing the total value of works completed. The format of the contractor’s statement must be approved by the supervision organization. A typical summary of the monthly statement is shown in Box 17.3.

Monthly
progress
payment

The value of the permanent works must be supported by detailed calculation showing measured quantity multiplied by the unit rate for each bill item.

Permanent
works
Variations

Under the FIDIC contract, the engineer is authorized to order the contractor to carry out extra works, omit any item of work, or make changes to the design of the project. If, for example, a culvert has been overlooked during project preparation, the engineer can order the contractor to construct it; if the embankment soil should turn out to have a lower than expected bearing capacity, the engineer can order an increase in the pavement thickness; and if a guardrail turns out to be superfluous, the engineer can order that it be deleted. The engineer can also instruct the contractor to use other materials than envisaged in the contract to produce a better or cheaper road. Variation orders must be given in writing, except when the actual quantities merely exceed or fall short of the quantities given in the bill of quantities due to inaccurate estimates during project preparation. For variation orders, it is often necessary to prepare separate bills of quantities. The value appearing in the summary of the monthly statement is the sum of all variations, each of which have been derived by multiplying the measured quantity of each item in the variation bill of quantities by the corresponding unit rate.

It is common in international works contracts to have a mechanism for adjustment of the contractor’s payments in the event of price fluctuations of materials, fuel, labour and the like. Such adjustments may be based on documented actual costs,

Price
adjustment

Box 17.3 Summary of contractor's monthly statement

<i>Contractor's monthly statement summary for period ending [date]</i>		<i>US\$</i>
1	Bill No. 1: General	
2	Bill No. 2: Earthworks	
3	Bill No. 3: Pavement	
4	Bill No. 4: Piling	
5	Bill No. 5: Structures	
6	Bill No. 6: Miscellaneous	
7	SUB-TOTAL to date	
8	Variations (+/-)	
9	Price adjustment (+/-)	
10	Dayworks (+)	
11	GROSS VALUE to date	
12	Retention (-)	
13	Materials on site (+)	
14	Mobilization advance (+)	
15	Repayment of mobilization advance (-)	
16	Liquidated damages (-)	
17	Interest for late payment (+)	
18	GROSS AMOUNT PAYABLE (GAP) to date	
19	GAP last statement (-)	
20	NET AMOUNT PAYABLE THIS STATEMENT	

which can be a cumbersome and unwieldy approach, or it may be based on a formula, an example of which is shown below.

$$\text{ADJ} = K + k_B \frac{B_1}{B_0} + k_C \frac{C_1}{C_0} + k_S \frac{S_1}{S_0} + k_T \frac{T_1}{T_0} + k_F \frac{F_1}{F_0} + k_L \frac{L_1}{L_0} - 1$$

where ADJ is the adjustment-factor for current payment certificate; K the fixed proportion of the contract price that is not to be adjusted for price fluctuations, for example, overheads and profit (e.g. $K = 0.39$); B the price index of bitumen; C the price index of cement; S the price index of steel; T the price index of timber; F the price index of fuel; L the price index of labour; 0 the subscript denoting price index at the time of tender; 1 the subscript denoting current price index; k_B, \dots, k_L the fixed proportions of the cost of the respective materials: bitumen (e.g. $k_B = 0.14$), cement (e.g. $k_C = 0.17$), etc.

The coefficients K and k_B, \dots, k_L have to be inserted by the employer in the tender documents based on the characteristics of the particular road project. The price indices are normally published by the government in monthly statistics.

Dayworks

The resident engineer may order the contractor to carry out minor work, for which there are no unit rates in the bill of quantities, as 'dayworks'. Such orders need to be in writing, and the works will be paid according to the rates given by the contractor

in the tender for equipment and labourers. If materials are needed, these will normally be paid at the purchase price (verified by receipts) plus an agreed mark-up (normally 30–40 per cent) to cover overheads and profit. Each day, the contractor submits to the supervision organization an account of the personnel, machines and materials utilized.

It is normal for the employer to withhold a certain amount of each progress payment, typically 10 per cent of the value of the works carried out. However, these deductions are discontinued or reduced when the total amount withheld reaches a certain proportion of the contract sum, often 5 per cent. Upon completion of the contract, the retention money is returned to the contractor.

To help the contractor's cash flow, it is normal to pay an advance amounting to 75 per cent of the value of materials brought to the site and not yet incorporated in the works. For the purpose of the contractor's monthly statement, a survey is carried out to determine the actual value of the materials on site at the end of the month in question. In this way, account is taken of increases caused by new consignments of materials and decreases due to materials being incorporated into the works. After the advance has been paid, the materials become the property of the employer.

The advance paid at the outset of the works will appear in full in each of the contractor's monthly statements. The mobilization advance is paid back progressively by the contractor according to rules laid down in the contract. At completion of the project, the whole advance will have been deducted through the monthly statements.

Liquidated damages are an amount of money to be paid by the contractor to the employer for losses incurred if the contractor fails to complete the project on time. The amount to be paid will be defined in the contract documents and is normally in the order of 0.05 per cent of the contract sum per day the contract period over-runs. In some countries, it is customary to set a ceiling on these deductions, such as 10 per cent of the contract value.

It is becoming increasingly common to compensate the contractor in the event of payments being made after the time limit laid down in the contract, for example, 45 days. This compensation would normally be tied to the country's bank lending interest rate.

The accumulated sum of all the above items will give the gross amount payable to the end of the period covered by the monthly statement. The net amount payable for the current month is found by subtracting the amount paid under the last statement.

The contractor should present monthly statements to the resident engineer as soon as possible after the end of the month in question. The supervision organization will then check the statement and issue a payment certificate within a given period of time, typically 30 days. The resident engineer is authorized to make corrections if there are errors or if certain works are not up to the specified standard, but is not allowed to withhold the payment certificate. After the payment certificate has been signed, it will be forwarded to the employer for payment.

Computer software has been developed for the preparation of monthly statements and payment certificates. This has greatly simplified the process. The inter-relationships between the various components regulating the payment due to the contractor are defined in the software. Each month, the quantities of works carried out are adjusted and fed into the software along with such variables as price

Retention
money

Advance on
materials

Mobilization
advance

Liquidated
damages

Interest for
late payments

Amount
payable

Payment
certificate

Software

adjustment indices and quantities of materials on site. Up-to-date progress tables and charts are then generated automatically.

Final payment Formal completion of the works is marked by the issue of a completion certificate. The final payment certificate can now be prepared. This document is a full and substantiated account of the project costs and shows the exact amount of money due to the contractor. Half of the retention money is normally released at this time.

Contract completion A ‘defect liability period’ is normally established, typically lasting one year. During this period, the contractor is liable to repair any defects that appear. At the end of the period, there is a final inspection and the ‘final acceptance certificate’ is issued, providing there are no outstanding defects that are the fault of the contractor. The remaining half of the retention money and the performance bond are now released, after which the contractor has no further obligation under the contract with the employer.

17.6 Progress control

Work programme The contractor must submit a work programme within a certain period of time after the engineer’s order to commence work. The FIDIC contract conditions have stringent requirements about the contractor’s obligation to produce and continuously update the work programme for the project. The planning of the works is the contractor’s responsibility, but the work programme needs to be approved by the resident engineer. All three parties to the contract have a vested interest in the work programme, as this is an essential tool to control the timely completion of the project. The programme also forms the basis for the employer’s cash-flow requirements and planning of the supervision activities. Finally, work programmes are necessary for the co-ordination with utility owners and other contractors.

Method statement The contractor is also required to submit detailed statements of the methods proposed for key items of work, complete with identification of plant and equipment. It is the duty of the resident engineer to approve or disapprove the method statement. If in doubt, the contractor should be requested to submit further information or clarification. The importance of a detailed and realistic work programme and method statement at the outset of the works cannot be overemphasized.

Programme assessment The resident engineer should assess the work programme to consider if there is a reasonable relationship between the time allocated for various work operations and the proposed input of machines and labour. Has the qualifications of local labour been taken into account? Are equipment output figures realistic? Has due consideration been given to local holidays and climatic conditions, for example, rainy seasons, when work efficiency will be reduced, if not halted? Has sufficient time been allowed for mobilization? Have all known obstructions and constraints, such as access to site, working arrangements on different sections of the site, the need for temporary bridges, late removal of buildings, etc., been incorporated into the work programme? The work programme should be returned for amendment where there are unsatisfactory answers to any of these questions. The contractor may have to improve the efficiency of his operations or mobilize more equipment and labour. Even approved programmes will have to be adjusted by the contractor, if work later falls behind schedule.

Types of programme Work programmes can be prepared in a number of ways. Common to all types is that the programme should be easy to read and easy to monitor. Simple bar charts,

where each activity is drawn as a bar on a time scale, are adequate for most road projects. As bar charts tend to become difficult to interpret if too many activities are shown, separate and more detailed charts can be prepared for each of the more important operations. A shortcoming of bar charts is that they are not well-suited to showing dependencies between work items, which means that it is difficult to assess the effects of localized delays. For larger and more complex work, including bridges, it is normally desirable to use network planning techniques and critical path analysis. Project management software is normally used to assist with this.

Many contractors, especially from developing and emerging countries, pay insufficient attention to work programming. For a programme to serve its purpose, progress needs to be monitored constantly. This means that actual achievement is compared with planned output. If and when work falls behind schedule, it is important to determine the reasons as soon as possible. Corrective measures, introduced at an early stage, can normally bring progress back on schedule. However, the real reason(s) for the delay must be clearly determined, if corrective measures are to be effective. It is not helpful simply to make changes to the programme itself, without addressing the reasons for the delay.

A key advantage of the FIDIC contract execution method is that it encourages projects to be completed within a fixed period of time. All too often, however, serious delays are incurred. The best way of preventing such delays is for the supervision organization to adhere closely to the FIDIC stipulations about work programmes and to work with the contractor to take timely corrective measures. To achieve this, the resident engineer and his staff may give advice, but must not interfere directly in the contractor's operations. Interference may result in subsequent cost claims.

Progress
monitoring

FIDIC
requirements

17.7 Extension of time

If delays have been incurred that are likely to cause delayed completion of the project, it will become a priority issue for the contractor to seek extension of time for completion. This is necessary to reduce or eliminate the liability to pay liquidated damages, which can be very expensive.

Under FIDIC contract conditions, extension of time is justified if, for example, delays have been incurred to the contractor's operation due to

- late availability of the site;
- late issue of drawings and instructions;
- suspension of work;
- unusually inclement weather;
- force majeure.

Two fundamental conditions must be met for time extension to be awarded. First, the delay must not be due to any fault of the contractor. Second, the delay must be such that it will directly influence the overall completion of the project, that is, it must fall on the critical path. An example of how to deal with a time extension claim is given in Box 17.4.

A request for extension of time must be made by the contractor as soon as possible after the delay has been incurred. The supervision organization is also under an

Need for time
extension

Time
extension
criteria

Procedure

Box 17.4 Example of determination of extension of time

The contractor claimed that earthwork operations within the road reserve had been stopped by landowners on two occasions. In the first incident, works were held up for 45 days while the employer sorted out matters. In the second incident, works were held up for 70 days. In consequence, the contractor requested $45 + 70 = 115$ days time extension.

The engineer agreed that this claim for time extension was admissible under the FIDIC contract, and that the incidents were not the fault of the contractor. It was furthermore agreed that there was a justifiable delay of 45 days for the first incident, as records confirmed that the team in question had been idle for the said period of time expecting that the problem would be resolved expeditiously. The second incident, however, was different because the resident engineer had instructed the contractor to move his team to another nearby location one day after the work had been stopped. The contractor had complied with the order and was able to return and complete the earthwork after the problems of the second incident had been settled.

The engineer noted that, as the contractor had four earthwork teams on the site, the effect on the overall production of one team being idle was only 25 per cent. Extension of time was therefore awarded as follows:

Incident 1:	Lost time		= 45 days
Incident 2:	Lost time	1 day	
	Removal of team	3 days	
	Re-entry	3 days	= 7 days
			<hr/>
Total			52 days
			<hr/>
	Actual delay (25 per cent of 52 days)		= 13 days

obligation to act fast, since the contractor will not be able to plan his operation properly until it is known whether or not an extension of time will be granted. When evaluating a request for time extension, the resident engineer is, as usual, under an obligation to be fair and reasonable. Granting additional time may result in increased project costs, as well as possible consequential cost claims from the contractor. On the other hand, rejection of justified requests may end up costing more, if taken to adjudication (discussed later). Extension of time should always be dealt with by the engineer, as knowledge of contract law is essential in handling these matters correctly.

17.8 Cost claims

Claim-prone contractors

Almost every road construction contract is faced with claims for extra payment. Sometimes the aggregate value of the claims equals or even exceeds the original contract sum. Some contractors deliberately tender low prices to secure the contract, and then start to prepare cost claims immediately work starts, with the help of claims

specialists, frequently lawyers. It is extremely important, therefore, that the resident engineer and his staff are conscious of claims from the outset of the project and make every effort to prevent or at least minimize extra costs. Some employers regard any request for extra payment as an ‘insult’ and expect the supervision organization to reject all claims out of hand. This is improper, since many claims are genuine and deal with matters that are risks borne appropriately by the employer under the FIDIC conditions.

The great majority of cost claims fall into one of the following three categories:

- The contractor has encountered conditions or obstructions that could not have reasonably been foreseen at the time of tender.
- Extra or changed works have been paid at too low unit rates fixed by the supervision organization.
- There have been delays and disruptions to work operations for reasons beyond the control of the contractor.

Types of claims

Unforeseeable conditions and obstructions often refer to matters that are hidden in the ground, such as problematic subsoil, poor quality borrow-sources, water pipes and the like. The basic rule is that the contractor is only entitled to recompense for any difficulties encountered if it can be established that even an experienced contractor could not reasonably have foreseen what happened. It should be noted that the FIDIC conditions require tenderers to make all such investigations as are necessary to prepare a realistic tender, including a survey of subsoil conditions. If the contractor has failed to do this adequately, then there will be no recompense.

Unforeseeable conditions

Additional or changed work is paid at the unit rates in the bill of quantities, if these are considered applicable by the supervision organization. In the event that existing unit rates are not applicable, or when no unit rates exist in the bill, the FIDIC conditions require that adjusted or new unit rates be agreed with the contractor. If it is not possible to reach an agreement on what is fair and reasonable, the supervision organization has the right to fix the rates unilaterally. In doing so, however, great caution should be exercised, as disagreements about fixed rates is one of the most common causes of dispute in international construction contracts. There are numerous examples where adjudications have resulted in very big awards to the contractor, and this has sometimes come as a surprise to the employer and the engineer. Therefore, variations should not be ordered without serious consideration of the possible implications.

Variations

Claims for delays and disruptions to the contractor’s operation are, in principle, admissible in the following cases:

Delays

- The site has not been made available on time.
- Drawings have been issued too late.
- The progress of the works has been formally suspended by the supervision organization.
- There have been errors and discrepancies in the contract documents.
- The survey information provided to the contractor has been in error.
- The contractor has been ordered to make excavations to uncover and test work which, subsequently, proved to be in order.

To qualify for payment, the contractor has to prove that it was not possible to utilize idle equipment and labourers, for example, by moving the team in question to alternative work sites. This is important, as it is the obligation of the contractor to make all possible efforts to minimize the effects of any obstruction to work. Furthermore, it is the contractor's duty to document all claimed expenses.

Case history

Box 17.5 shows the importance of evaluating claims in strict accordance with the conditions of contract. The two examples appear quite similar. Both deal with the late issue of drawings, but the outcomes are nevertheless very different.

Misconceptions

The following points address commonly held misconceptions about cost claims:

- The fact that a time extension has been granted does not automatically justify additional payment to the contractor.
- Exceptional weather may entitle the contractor to time extension, but not to extra costs.
- Whether or not the contractor has lost money is irrelevant and cannot form the basis of a cost claim.
- Any misdeed of a subcontractor is the responsibility of the contractor.
- As long as the resident engineer and his staff act within the framework of the contract, there is no such thing as 'too strict' supervision.

Box 17.5 Examples of claims for delay

Example 1

One culvert had inadvertently not been shown on the contract drawings. When the resident engineer discovered this omission and ordered a variation, the culvert works were nearly completed. The contractor immediately ordered the additional culvert pipes, but one work team, which otherwise would have been demobilized, had to remain idle for three weeks waiting for the pipes to arrive on site. The contractor claimed that compensation should be paid for the additional costs incurred. The engineer accepted this claim as

- it was admissible under the contract;
- there was no alternative work for the team in question;
- the contractor could not reasonably have foreseen the need for this culvert.

Example 2

One pipe culvert was shown on the drawings, but the required dimensions had inadvertently been left out. As for *Example 1*, culvert works were nearly completed when the resident engineer discovered this error and provided the contractor with the correct dimensions. As before, one culvert team was forced to be idle for three weeks, until the pipes of the specified dimension arrived. The engineer rejected this claim on the ground that it was the contractor's contractual obligation to study the drawings at the time of the tender. If this had been done, the error would have been discovered and clarification could have been sought. It was held by the engineer, therefore, that the contractor could have prevented the culvert team becoming idle.

When raising a cost claim, the contractor has to comply with strictly formal procedures. First, it is important that written notice is given about the intention to claim extra payment. Although the claim itself may be submitted later, it will not be disqualified if the requirement for a specific notice period has been complied with. Second, for a claim to be valid, it must be raised with reference to a specific clause (or clauses) in the contract. Finally, the burden of proof is on the contractor, who needs to provide all documentation to substantiate the additional costs claimed.

Claims
procedures

It is normal for the engineer to retain the authority to deal with cost claims. It is the intention of the FIDIC conditions that the engineer acts as an independent mediator and makes fair and unbiased decisions, even if decisions might be unfavourable to the employer.

Mediation

If the contractor does not accept the engineer's decision, the engineer should be informed accordingly in writing. If a dispute cannot be resolved through negotiations, it will be referred to adjudication, where it will be settled by one or more adjudicator, appointed jointly by the two parties. In case the two parties cannot agree on the appointment of an adjudicator, each party may apply to any appointing authority named in the contract or, if none, to the President of FIDIC, to appoint an adjudicator. Such appointment is final and conclusive. In an adjudication, it is crucial that the supervision organization is able to present written records of instructions given to the contractor, observations, tests, progress of work, the weather, equipment on site, personnel, material supply, errors by the contractor, etc. The engineer will be in default of obligations under the contract if such evidence cannot be produced. The adjudicator's decision is final and binding.

Adjudication

17.9 Default of contractor

According to the FIDIC conditions, the employer can expel the contractor from the site, if the engineer certifies that the contractor has

Expulsion of
contractor

- not commenced the work or has abandoned the site before completion of the works, or
- assigned the works or parts thereof to others without permission, or
- failed to correct faulty work which has been rejected by the supervision organization, or
- become bankrupt, or
- flagrantly and persistently failed to fulfil obligations under the contract.

In the event of an expulsion, the supervision organization will ascertain the value of all completed and started work, as well as the value of all equipment, plant, materials, scaffolding and the like. The employer can either complete the project with its own organization, or can select a new contractor to carry out the remaining works. In either case, at the discretion of the employer, the equipment, materials and other site assets of the original contractor can be sold or utilized in completing the project. The fact that the contractor has been expelled does not mean that the contract with the employer has been rescinded, and the contractor is still liable for the costs of completing the works.

Completion of
works

Settlement of account

The project account can be finalized only after the project has been completed by others, and is based on the following principles. First, the supervision organization will determine the payment that would have been due to the original contractor, if they had completed the works. From this amount is deducted the actual cost of completing the project, plus liquidated damages and any additional costs incurred by the employer as a result of the default. In consequence, the original contractor will inevitably end up owing the employer a large sum of money. To recover this debt, at last in part, the employer can confiscate the performance bond and retain any proceeds from the sale of the equipment, etc. If, as is normally the case, there is still a debt due from the contractor, the employer's only option is to try to recover this through civil litigation.

17.10 Supervision procedures

Meetings

Meetings are an essential tool in project administration. To be successful, meetings need to be well prepared, business-like and conducted in accordance with a fixed agenda. It is normal for the resident engineer to take the chair. Clear and concise minutes of meetings need to be taken for subsequent circulation and approval of all participants.

First site meeting

Prior to commencement of work, it is normal to hold a site meeting with participation of the engineer, the resident engineer and other supervision staff on the one side, and the contractor's project manager (site agent), other representatives and subcontractors, on the other side. It may be a good idea to invite the employer to this important first meeting, which is held for the following main purposes:

- to acquaint the parties with each other, including an introduction to the contractor's and the engineer's organizations;
- to clarify which powers and authorities have been delegated to the resident engineer;
- to agree administrative routines, including lines of communication and organization of day-to-day activities;
- to review the contractor's work programme and establish how far advanced is mobilization;
- to verify that contract securities and insurances are in place;
- to agree how site sections are handled, if relevant;
- to explain all pertinent issues to the contractor, including relations with local property owners, utility authorities and the like.

At the first site meeting, the supervision organization will also seek to clarify and settle any doubts the contractor might have about efficient future co-operation.

Routine site meetings

During the works, it is good practice to hold regular site meetings for the purpose of recording events, discussing problems and taking decisions. The frequency of these meetings varies, but monthly meetings are the accepted norm on many internationally tendered road projects. The main participants are the resident engineer and the contractor's project manager, both supported by key technical and administrative staff. The agenda, which is circulated prior to the meeting, will often

comprise the following topics:

- Matters arising from the last site meeting
- Progress in the previous month
- Planned progress
- Personnel, machines and materials
- Technical matters
- Financial matters
- Administrative matters
- Miscellaneous
- Time for next site meeting.

It is the principal duty of the resident engineer to keep records of whatever transpires on the project. These records form the basis for approval of works as well as payments to the contractor. Furthermore, site records are indispensable in the event of disagreements; for example, in respect of extra or changed work, time extensions and claims for additional payment. Supervision records may broadly be divided into the following four categories:

Records

- Historical records; that is, work programmes and monitoring data, weather data, resident engineer's diary and daily inspection records.
- Quality records; that is, test results, survey control, etc.
- Quantity records; that is, measurements for payment, monthly statements, payment certificates and variation orders.
- 'As built' records; that is, drawings and descriptions of all completed parts of the project.

Correspondence to and from the contractor, the employer and various other parties form part of the resident engineer's records. Because of the sheer volume of paperwork involved, it is important that an adequate filing system is devised and adhered to from the outset of the construction.

The supervision organization is normally required to prepare monthly progress reports for distribution to the employer, the international financing agency and other parties with an interest in the project. Typically, this report will have the following contents:

Progress reports

- A brief description of the original project and any significant changes.
- Key contract data, including any approved time extension.
- Information about approved subcontracts.
- Progress of work compared to the approved programme; identification of matters that may cause delays, and action taken to bring the project back on schedule.
- Physical considerations, including data on weather, construction resources and testing, as well as quality issues.
- Financial considerations, including the last month's payment certificate, schedule of variation orders, cost claims approved and pending, and updated cash-flow projections.
- Supervision organization, performance and problems encountered.

A good resident engineer will take pride in preparing informative and lucid monthly reports. The use of progress photographs is a good idea, as it is well known that one photograph can tell more than many words.

References

- FIDIC (1989) *Guide to the Use of FIDIC Conditions of Contract for Works of Civil Engineering Construction*, 4th edition, Lausanne: Fédération Internationale des Ingenieurs-Conseils.
- FIDIC (1999a) *Conditions of Contract for Construction*, Lausanne: Fédération Internationale des Ingenieurs-Conseils.
- FIDIC (1999b) *Short Form of Contract*, Lausanne: Fédération Internationale des Ingenieurs-Conseils.

Appropriate technology

Jan de Veen, Richard Robinson and Bent Thagesen

18.1 Technology options

Road works, whether construction or maintenance, can be carried out through a wide range of methods using various mixes of labour and equipment. These methods can be characterized by

Concepts and definitions

- *Labour-intensive* – use of labour for all activities, including only unpowered hand tools, and walking to site from home or road camp.
- *Intermediate or labour-based* – labour intensive, as above, but supplemented by tractor-based transport and simple equipment, and transport to site where appropriate.
- *Equipment-based* – predominant use of equipment with high output (sometimes called ‘equipment-intensive’ or ‘capital-intensive’).

The aim should be to use the most suitable mixture of labour and equipment in a given social, technical and economic context. This is known as ‘appropriate technology’.

18.2 Choice of technology

18.2.1 Suitability of technology

Most investment in the transport sector has traditionally been spent in a capital-intensive manner. However, the development of road transport infrastructure has been recognized as a sector where it is very appropriate to test the use of different technology options. More than 40 developing countries have now implemented civil works projects using labour-based methods (ILO 2000). To develop technological alternatives that are locally sustainable and deliver technically high quality products, modern management methods and a long-term approach have to be adopted. It is necessary to develop or adapt policies, implementation strategies and legislation. Equally, investments have to be made in improving the business and management skills of local entrepreneurs and in the development of a working environment in which they can compete successfully.

Enabling environment for labour-based methods

- Criteria When deciding on the most appropriate technology in a given situation, the following criteria need to be considered:
- The ability of the technology to meet the designated standards, recognizing that these may vary depending upon
 - road class and traffic level;
 - geographic and climatic conditions.
 - The cost-effectiveness of the approach, recognizing that this may depend upon
 - the relative cost of labour and equipment;
 - the procedures adopted and the effectiveness of management of the operation.
 - The availability of both labour and equipment, recognizing that this may depend upon the demographic and socio-economic conditions in different parts of the country for labour, and upon the type of equipment being considered.
 - A preference for many countries to use domestic resources, rather than those that are imported.

Technology options for road works are reviewed here against these four criteria.

18.2.2 Technological appropriateness

Existing equipment-based practice Equipment-based road construction and maintenance methods developed in the industrialized countries are used extensively in developing and emerging countries. There are several reasons for this:

- Politicians, planners and engineers are strongly influenced by the technology used in industrialized countries, and there is a very effective lobby and sales pressure from heavy equipment manufacturers.
- Most international or locally established road contractors possess heavy equipment and are used to equipment-based working methods; they have little experience of the organizational and managerial requirements of working with people instead of machines.
- Heavy equipment is often supplied at no apparent cost to government through foreign funded aid packages; tax exemptions can also distort the real cost of machine-based methods; some local contractors ignore the replacement costs of the machines when pricing works.
- Employment creation and local skill development are seldom taken into account when decisions are made on project implementation.
- Lending criteria of financial institutions generally favour large-scale programmes that have higher disbursement levels than projects executed with labour-based methods.

Factors determining appropriateness The appropriateness of a given technology option in a particular country will depend principally upon the (Edmonds and de Veen 1991)

- level of economic, socio-economic circumstances, such as wage levels, the level of unemployment and underemployment;

- existence of options in the work locality, in terms of labour availability, climatic and terrain conditions.

In countries with no previous experience with labour-based road works, there are likely to be many obstacles to the introduction of this technology. These relate to the use of inappropriate procedures, preconceived ideas and attitudes, rather than the technical feasibility of the technology. Options for technology need to be developed before a real choice can be made. It is essential to make the different levels of government and administration aware of the socio-economic implications of not utilizing local resources, and to demonstrate that labour-based approaches can work effectively in individual countries, to facilitate the inception and implementation of labour-based programmes. Demonstration projects with technical assistance will generally be an indispensable first step to provide a sound base for larger-scale programmes (de Veen 1991). These can be a valuable method of increasing awareness, within governments and road administrations, of the benefits of labour-based programmes, and of the socio-economic impacts of failing to maximize the use of local resources.

Demonstration projects

18.2.3 Cost-effectiveness

The costs of alternative approaches should underpin decisions on the most appropriate choice of technology. These should take into account costs and productivity of labour and equipment-based approaches including, for each alternative, the costs of staff, material, tools and equipment, camps, offices, operating expenses and overheads. Figure 18.1 shows factors to be considered when making cost comparisons. In addition, the cost of equipment replacement, mobilization, headquarters staff and training facilities should not be forgotten.

Cost estimates

An important factor in the assessment of whether labour-based techniques are economically appropriate is the daily wage rate. The agricultural wage rate is a rough indicator that can be used when establishing wage rates for road works in rural areas. Research has indicated that labour-based methods are competitive when wage rates are less than US\$4 per day (Stock and de Veen 1996).

Labour costs

Labour costs will depend on productivity. Labour productivity data may be taken from international sources, if no systematically organized labour-based work has been undertaken recently in the country. Such data should be used only for tentative planning and the aim should be to establish local data as soon as possible. Daily productivity can be assessed relatively quickly through work studies. However, only experience over a long time can indicate the likely frequency of work disruptions and, hence, the long-term productivity.

Labour productivity

Productivity data for equipment may be taken from manufacturers' handbooks. Allowance should be made for the fact that data provided by the manufacturer tend to be optimistic and reflect the conditions in the country of manufacture, normally an industrialized country. Equipment productivity in developing and emerging countries should reflect the likely low utilization caused by poor equipment management and planning, and high non-availability owing to breakdowns and shortage of fuel and spare parts. Figure 18.2 shows the factors that influence the productivity of labour and equipment.

Equipment productivity

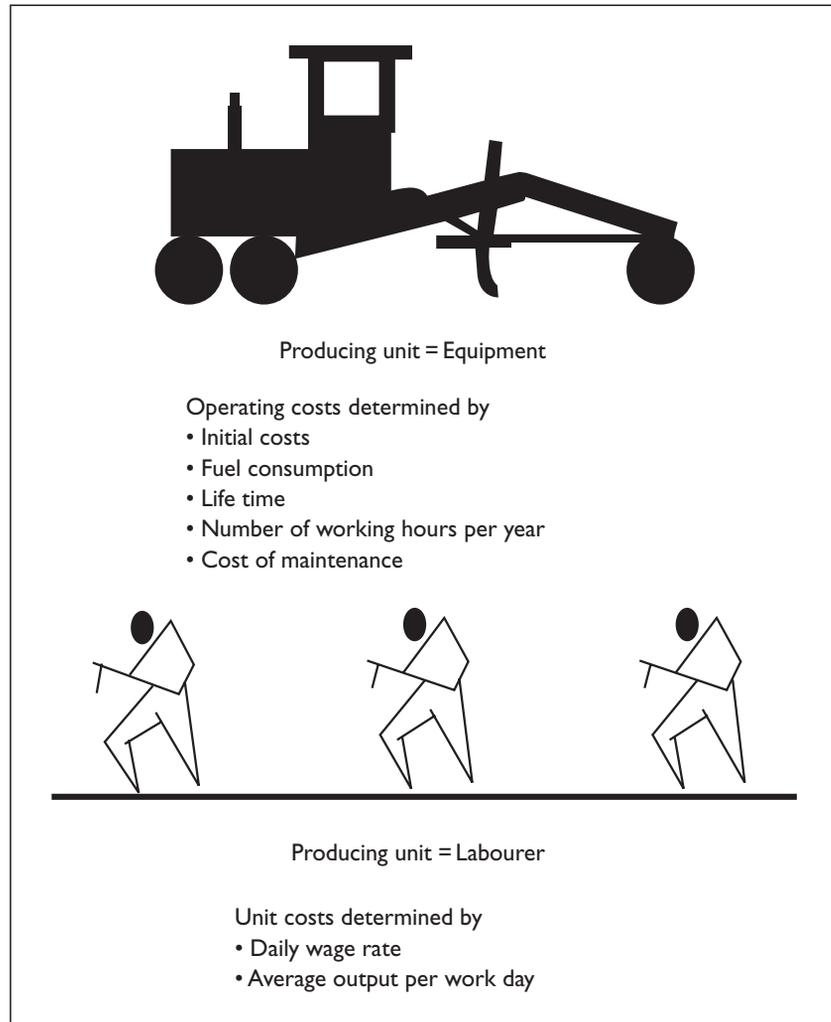


Figure 18.1 Comparison of cost between equipment and labour.

Copyright: International Labour Organization (Beusch and de Veen 1991).

18.2.4 Availability of labour and equipment

Labour availability

The availability of labour, in sufficient numbers, is a prerequisite for any labour-based works or maintenance programme. Normally, the level of unemployment or underemployment is used as an indicator of labour availability. Unfortunately, these figures are difficult to obtain, in particular in relation to the non-wage economy in rural areas. Population density can be used as a rough proxy for labour availability, but these figures must be treated with care as a low value does not necessarily mean that labour is not available in a particular region. Also, experience has shown that, even in very sparsely populated regions, the availability of paid work has attracted many workers from surrounding areas. On the other hand, seasonal and well-paid work opportunities – such as coffee picking – may draw off labour from road works even in highly populated areas. A population density of about 25 persons per square kilometre provides a threshold

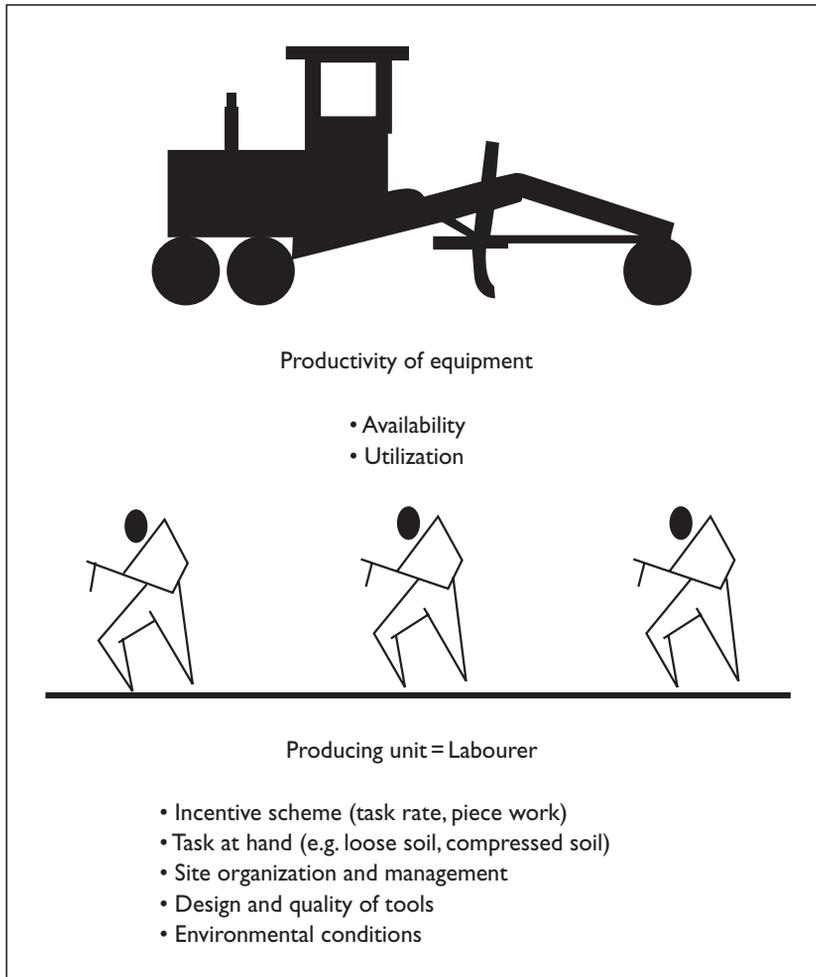


Figure 18.2 Comparison of productivity between equipment and labour.

Copyright: International Labour Organization (Beusch and de Veen 1991).

above which labour supply should not normally be a problem. Labour availability studies should be undertaken for areas where the population density is lower.

Attitudes of the available labour and local customs make it necessary to ascertain the willingness of available labour to accept employment on road works. For example, earthworks is a predominant activity in most road construction projects. Yet, in certain countries, earthworks activities rank below agricultural work in the hierarchy of employment. In other countries, it is not acceptable for women to carry out certain types of road works.

The Chinese and Indian experience demonstrates that labour-based methods can be used cost-effectively even for significant earthworks in large schemes. However, labour-based methods are best suited for programmes comprising a large number of small, technically simple and geographically dispersed activities. Works involving the movement of large quantities of materials over large distances, or substantial production in stony soils or rock, is best carried out by mechanical equipment. Paving with

Labour
attitude

Project type
and location

pre-mixed asphalt requires equipment if high quality is required. For gravel road construction, all or nearly all activities can be carried out by labour. Also labour-based methods are well suited to routine and periodic maintenance works.

18.2.5 Domestic resources

Affordability of labour and equipment

In many developing and emerging countries, increasing debt-servicing obligations and less income from exports have had severe impacts on national economies, forcing countries to restrict imports and to devalue local currencies. Foreign exchange shortages have often resulted in foreign goods becoming unobtainable or extremely expensive. This has had a significant effect on operations in the road sector, with its traditional dependence on foreign equipment and goods. At the same time, low-skilled employment opportunities, particularly in the agricultural and industrial sectors, are on the decline, and this is exacerbated by high population growth. As a result, unemployment and underemployment are rising in a large number of countries and wage levels are decreasing in real terms. Against this background, employment-intensive alternatives of carrying out road programmes are often attractive.

Promotion of domestic resources

Many countries wish to promote the use of local resources (human as well as material) for infrastructure programmes. This can involve, on the one hand, offering employment opportunities to the local population and, on the other hand, the development and involvement of the indigenous works industry (local contractors, consultants and manufacturers). Such approaches can also maximize the use of local materials. However, preconceived ideas and attitudes, particularly of government engineers, can impede the introduction of technology options based on the use of domestic resources. Exposure of both politicians and implementers to well-managed labour-based projects through study visits has proved to be very effective in changing such attitudes. Demonstration projects can also be effective.

Demonstration projects

A demonstration project can often be helpful to prove the viability of a given approach, and to define and introduce the systems and procedures. Once the technical and financial viability of a particular approach has been proved, a transition period of several years is usually necessary to enable the establishment of a large-scale and long-term programme. During this period, procedures will need to be standardized, training programmes installed, and administrative and organizational systems refined and adapted to the different requirements of large-scale programmes.

Shadow prices

'Shadow' prices should be used rather than market prices when taking decisions about choice of technology to reflect the real economic costs and implications of applying a given technological approach. Shadow prices reflect the opportunity foregone of using the resources in question for other purposes (see Chapter 7). The socio-economic advantages of using labour-based methods (employment, local skills development, spin-offs for the rural economy) usually mean that labour-based technologies can be priced at a lower level than actual financial costs. On the other hand, the disadvantages of using equipment-based methods (use of foreign exchange, often poor sustainability) can lead to a higher pricing than actual financial costs.

Skill development

Domestic resources need to be developed so that they can be applied cost-effectively, but without compromising on quality and output levels. However, special skills are required for organizing, administering and controlling large labour forces. Although these skills are available even in remote rural areas of developing and

emerging countries (secondary school leavers, artisans, local contractors and consultants), they need to be encouraged and developed. Reasonable production and productivity levels can only be obtained by the use of appropriately trained management and technical staff. Training investments may need to be made.

18.2.6 Steps in technology choice

In general, the steps to be taken in choosing technologies for all types of road works in developing countries are indicated in Figure 18.3.

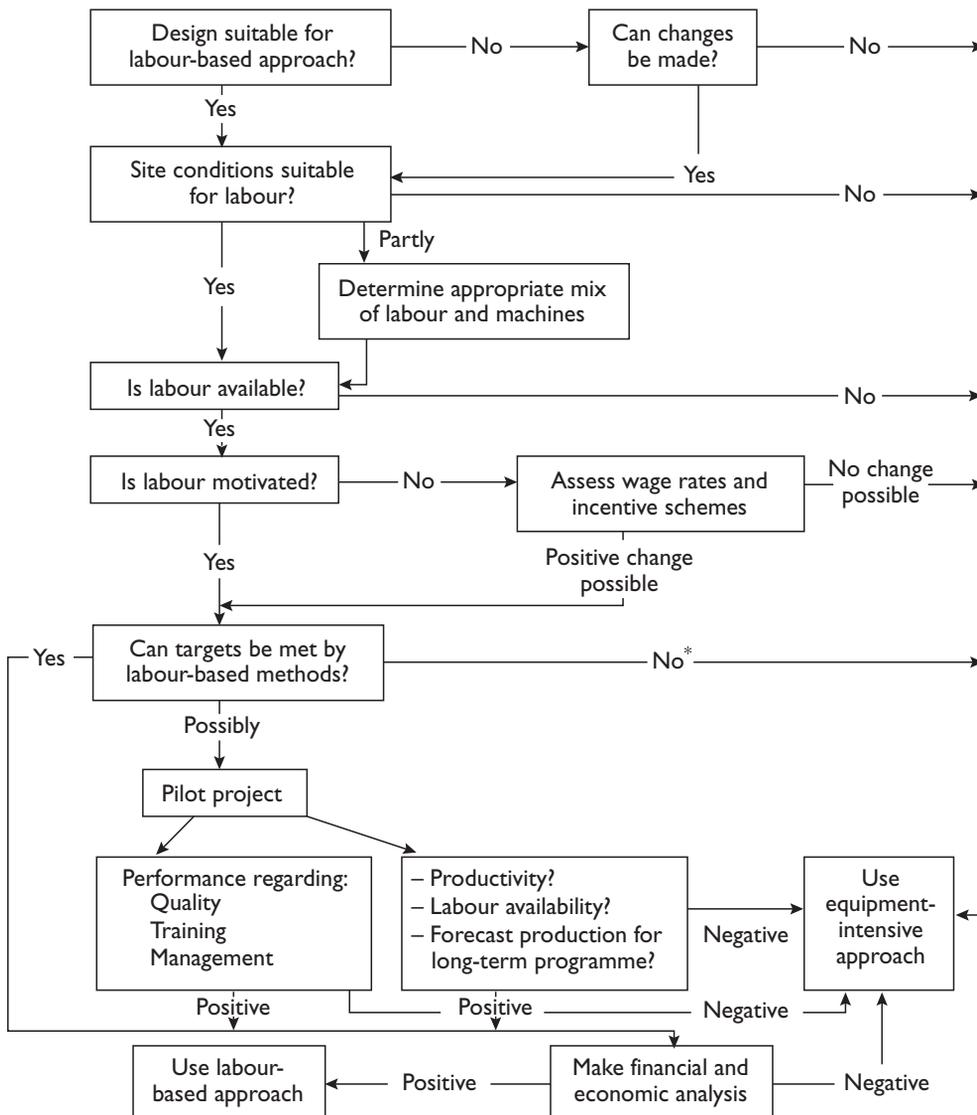


Figure 18.3 Decision-making regarding technology choice.

Copyright: International Labour Organization (Beusch and de Veen 1991).

Note:

* Based on actual performance/comparison between labour-based and equipment-intensive alternatives.

18.3 Intermediate methods

18.3.1 Institutional framework

Institutional context Labour-based road works may be undertaken by a government agency using its own resources ('in-house works', 'direct labour' or 'force account') or, alternatively, using private contractors. In either case, it is important to define clearly the institutional framework in which labour-based projects will operate. Often this is not a simple matter because different agencies may be responsible for the planning, construction and maintenance of different classes of roads.

Organizational responsibility In some cases, the responsibility for labour-based road projects has been placed with local government rural development agencies, whose principal objectives are to employ the rural poor and to develop local organizations for community projects. In other cases, an operational division within the road administration devoted exclusively to labour-based infrastructure works has proven to be very effective for larger scale operations (de Veen 1980). It is generally better to make use of existing institutions to meet new needs rather than to create new institutions. At the national level, the responsible institution will need to ensure the following:

- development of project selection criteria;
- preparation of budget requests at central level;
- development of effective payment, planning and reporting procedures;
- preparation for maintenance activities;
- disbursement of funds to lower levels as appropriate;
- decentralization of the planning, execution and control functions of labour-based projects to the greatest extent possible;
- encouragement of local participation in the decision-making;
- the framework for the use of self-help and other types of contribution by the local population.

Maintenance considerations Future maintenance requirements and methods should be considered when the choice of works technology is made. Often, many road maintenance activities can be carried out by locally employed people. In these cases, future maintenance workers can be selected from the best construction workers (Jones and Petts 1991).

18.3.2 Contractors

Advantages Executing road work by competitive contract usually has a number of advantages, including lower costs. Contractors tend to be more flexible. They can motivate their staff in a variety of ways that would be difficult for government agencies. They are more likely to be focused on getting the job done, with the likelihood of less delay. Furthermore, their knowledge of the local environment in which they work often enables them to utilize locally available resources in an imaginative and optimum way. However, contracts let without competition result in high costs due to poor management or excessive profit taking. Maintaining a small in-house account capacity to set standards, and to carry out emergency works and maintenance can be helpful in resisting the effect of contractor cartels.

Local contractors Large contractors tend to be biased towards using equipment-based methods. Small, local contractors are likely to have little equipment and, therefore, be more

interested in undertaking labour-based works. In a number of countries, programmes are being established or are ongoing, which work entirely through local contractors. This type of programme has the multiple objectives of introducing cost-effective labour-based methods and developing the private sector. However, small contractors will generally need support to enable them to develop the ability to price, bid, and to administer their businesses and their operations. Cash-flow problems and poor administration have often caused the bankruptcy of technically competent contractors. Training in business management and work methods must be provided to develop the required organizational and management skills.

Subcontractors can be used by the main contractor to provide labour or materials, or to undertake specific items of work, such as culverts and bridges. Subcontracting can be a useful mechanism for developing independent contractors. Training and assistance may be needed for this, and any project design should then make it attractive to the main contractor to provide such support (Bentall *et al.* 1999).

Contract procedures for road works are generally developed for the execution of large-sized projects by equipment. Standard specifications, conditions of contract, and methods of tendering all tend to reflect this. If local firms are to participate in contracts for small jobs, contract documentation needs to be simple and straight-forward. At the same time, the bills of quantities and the work descriptions need to be sufficiently detailed to guide the small contractors in their estimating. The very size of contracts may also prevent local contractors from bidding. If a contract is broken down to a number of smaller-sized elements, the work will be more accessible to local firms.

To be able to tender for road works, contractors are usually required to be registered. To do this, contractors must normally possess a minimum amount of equipment. For small labour-based contractors, it should be made possible to use a proven management experience or a training certificate as a pre-qualification requirement. Countries, such as Ghana, have created a separate category for labour-based contractors, allowing them to bid for works of a certain nature and up to specified amounts.

The organizational transition from in-house to contract work is difficult to make. The supervising agency should develop 'in-house' skills to enable the award, administration, supervision and control of the works carried out by contract. There may still often be scope for employing a local consultant, who will act on the client's behalf in respect of supervision and control of the works. System development for contract assessment, award and control, and training will normally be indispensable to enable the client and consultants to carry out their roles effectively (ILO 2000). In most cases, it will be necessary to develop contract procedures and a payment system appropriate to small local contractors.

18.3.3 Labour issues

The greater part of the works in labour-based road projects is carried out by temporarily employed local people. These people do not become a part of the permanent labour force, and their conditions of work are fundamentally different. Even so, labour laws concerning minimum wage, minimum age, non-discrimination, elimination of forced labour, workers' compensation for work accidents, and safety and health should be respected (Tajgman and de Veen 1998). However, existing labour legislation and regulations generally do not fit a situation where temporary workers are employed on a large scale. Stakeholders, such as project managers,

Subcontractors

Contract procedures

Contractor registration

Developing contractor management capability

Legislation and regulations

contractors' associations, workers' organizations, or government ministries, should initiate a discussion on the adaptation of the legislation.

**Basic
legislative
requirements**

Often new legislation, relating to temporary workers, will need to be established. This should cover issues such as recruitment and dismissal procedures, workers' rights to associate, motivation and discipline, and safety and health, including minimum facilities, such as drinking water, latrines and protective clothing, to be provided on site. The legislation should also cover issues such as (partial) payment in food or in kind, entitlement to paid leave, workers' rights in case of illness or accident, duration of the employment, etc. As far as possible, legislation should be put in place by the stakeholders in consultation with the Ministry of Labour before the start of a project.

Contracts

For road works carried out by contract, the relevant legislation and regulations should be reflected in contract documents. Conditions of contract and contract clauses should clearly spell out the rights and obligations both of the contractor and the worker, as well as the sanctions if these are not adhered to.

**Career
development**

As equipment-intensive methods in many countries have been the prevalent way of constructing roads, staff regulations are normally geared towards their use, with emphasis on academic qualifications. For labour-based work, supervisors need to have people-management and organizational abilities rather than sophisticated training. This does not mean that their work carries little responsibility. Unfortunately, government regulations often do not recognize that individuals with little formal education can take on this kind of responsibility and, therefore, their employment and promotion to levels corresponding to well-educated staff can be difficult to accomplish. Institutionalizing labour-based work requires that this issue is addressed. This may require the recognition of labour-based training certificates by the Public Service Commission, or the lowering of recruitment standards for labour-based supervisors. Career development possibilities of trained staff should be ensured through the integration of labour-based training courses into the overall training programme of the road administration. These courses should be given the same status as other courses and, once successfully completed, they should become a stepping stone for promotion to a higher rank in the organization.

18.3.4 Technical issues

18.3.4.1 Earthworks

**Design
requirement**

It has proven to be feasible and economical to execute even large earthworks by labour, provided that there is not too much hard rock, and provided that haul distances are minimized. Haul distance can be reduced by moving soil sideways rather than longitudinally, with the need to balance cuts and fills along the length of the road, which is usually the case when heavy equipment is used. Since high design speeds are not always required for low-volume roads, road alignment can follow the terrain contours more easily, reducing the earthworks in these cases. Labour-based methods benefit from a different approach to design than when equipment-intensive means are used. In the design, attention should be paid to the method of working. For example, the ditch cross-section should depend on the technique used. Blade graders can only produce triangular V-shaped ditches. Ditches excavated by shovel should have a flat bottom.

Large mechanical compaction equipment is difficult to move between the many sites being worked by labour crews on large works projects, and is usually underutilized on smaller projects. It is more appropriate to use light compaction equipment on labour-based projects. In some cases, indirect or natural compaction can be used, such as consolidation resulting from the action over time of the weather and traffic on an embankment. The main disadvantage of indirect compaction is the level of erosion and deformation during and immediately following the works period. Extra work-days are required to reshape the road one or more times before the required compaction standard has been achieved. Also, the design specifications in contract documents will often specify thick layers of base and sub-base materials, coupled with high compactive effort: this assumes construction by machine. If thinner layers are used, it may be possible either to reduce compaction standards, or to use materials of a lower quality. Using lower quality materials may also reduce haul distances. In Mexico, loaded trucks or tractors have been successfully used to achieve initial compaction. Pedestrian-controlled vibratory compactors are used in Kenya.

Compaction

18.3.4.2 Pavements

Some pavement types are more suitable for labour-based methods than others (World Bank 1983).

Natural gravels are the most economical to use as they can be laid without any double handling (see Figure 18.4). It is easier for labourers to lay fine gravels than gravels containing large stones if a good surface finish is required. Labour-based methods are not able to achieve the same high standard of surface finish as a skilled grader driver, but spreading gangs can build up considerable skill if their work is well supervised and checked frequently with string lines.

Gravel

Hand mixing of lime-stabilized soil is used, for example, in China. However, hand mixing is inefficient for pulverizing clayey materials and lime dust may create health problems for the labourers. Hand mixing of cement-stabilized soil is not recommended as cement reacts and sets too quickly.

Soil stabilization

Surface dressing may be carried out by labour-based methods and the quality of work can be almost as good as by mechanized methods, if the works are properly managed. Material specifications are the same as for mechanized surface dressing and described in Chapter 14.

Surface dressing

Hot premixed asphalt are sometimes produced by hand. However, it is difficult to create satisfactory working environment for the labourers when the bitumen and the aggregate are heated over an open fire. It is not possible to obtain the same quality of mix and the same standard of surface finish as by the use of mechanized mixing plants and asphalt pavers. However, hot-mixed asphalt, mixed and spread by hand are used on many thousands of kilometres of roads throughout India and China. In many other countries, this type of asphalt is widely used for repair works. Cold-mixed asphalt, consisting of aggregate and cold emulsified bitumen, is better suited to hand mixing. Troublesome smoke is avoided and correct proportioning is easier. The use of cold-mixed asphalt should be encouraged.

Premixed asphalt

Penetration macadam consist of two or three layers of progressively smaller, angular stones. Each layer is compacted by rolling, after which it is sprayed with hot bitumen, cutback bitumen or emulsified bitumen. It is well-suited to labour-based

Penetration macadam



Figure 18.4 Gravelling rural access roads in Kenya. Photograph courtesy of TRL Ltd.

methods. Penetration macadam is used extensively in Indonesia. Stones for penetration macadam may be broken by hand, but labour is generally only competitive for size of stone above 25 mm.

18.3.4.3 Working methods

The use of correct working methods is crucial for labour-based work to ensure a high labour productivity. Appropriate working methods are described in different handbooks and training materials (Antoniou *et al.* 1990; Beusch and de Veen 1991; Karlsson and de Veen 1992). Figure 18.5 is an example, illustrating how soil from the ditches should be spread to give the correct camber on earth roads.

18.3.5 Tools and equipment

Requirements The quality and design of hand tools and light equipment have a significant influence on worker productivity and, ultimately, on the cost and speed of work. The timely availability of sufficient quantities of well-designed, good quality tools should be ensured for the different activities and conditions. Good tools should

- perform effectively the function for which they are intended;
- be strong and durable enough to have an acceptable working life, bearing in mind normal abuse;
- be correctly proportioned to the body dimensions of the user;

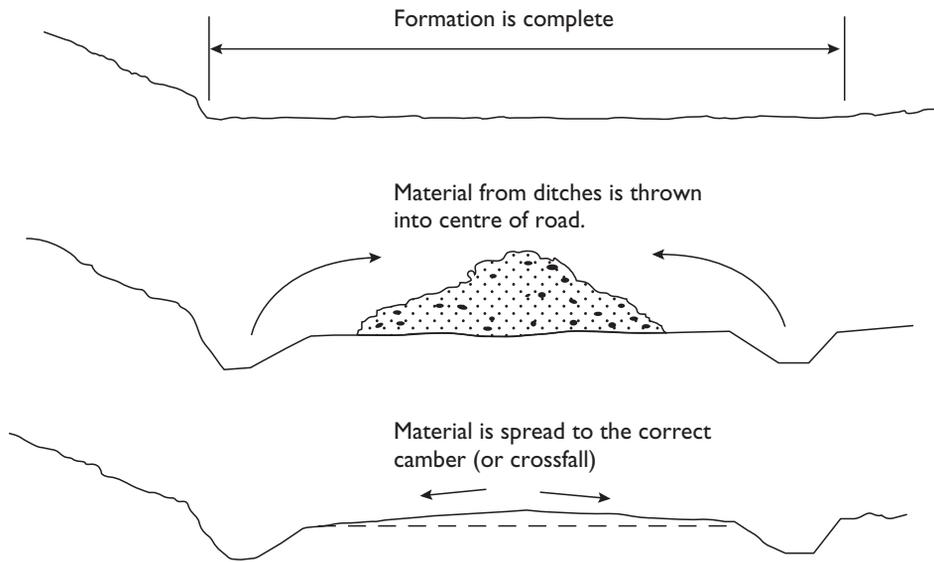


Figure 18.5 Digging of ditches.

Copyright: International Labour Organization (Antoniou *et al.* 1990).

Notes

- 1 Surface drainage prevents water from damaging the road by leading it off the road quickly. This is done by shaping the road so that the water can flow freely into the side drains.
- 2 The slope from either side of the centreline to the sides is called the camber. This kind of sloping is used in open terrain, with a ditch on each side. In sidelong ground, with only ditch, it is better to use a crossfall.

- match the strength and working capacity of the user;
- be safe in use;
- have as low as possible initial cost.

To ensure that tools and handles meet the quality and design requirements of heavy-duty road works, specifications for them should be explicit and detailed. At the purchasing stage, standard testing procedures should be applied that allow for acceptance or rejection of tools on an objective basis. The *Guide to Tools and Equipment* (ILO 1981) provides optimum designs and quality standards, and describes appropriate test procedures for a wide range of hand tools and light equipment. In most countries these principles are not applied. Tender board procedures are geared to the evaluation of major expensive items rather than to large quantities of hand tools and light equipment. The practice of purchasing at lowest prices, without giving weight to technical and quality aspects, is the main reason for the presence of poorly designed, bad quality tools in labour-based projects.

Specifications

Earthmoving is generally the most expensive works operation. There is a close correlation between the preferred equipment and the length of the haul. Appropriate equipment includes

Earthmoving

- head baskets
- wheelbarrows

- animal carts
- tractor–trailers
- two-wheeled tractors
- trucks.

Head baskets

The head basket is simple and very cheap for short hauls of 30 m or less, particularly in difficult terrain (see Figure 18.6). Workers need to be accustomed to its use.

Wheelbarrows

The wheelbarrow is probably the single most useful item of equipment for labour-based road works. For haul lengths greater than about 20–30 m, their use increases the productivity of labour compared with head loading. However, many of the wheelbarrows commonly in use are of an inefficient design and are poorly manufactured. The weight distribution is wrong, and the size and type of wheel is ill-suited to works sites. A more appropriate design is illustrated in Figure 18.7.

Animal carts

Animals can be a very appropriate source of power for haulage. However, in the countries where animal carts are common, there is considerable scope for improving their general design. The major drawbacks of 'traditional' vehicles are excessive unladen weight, poor tyres (with frequent punctures), overloaded bearings which are expensive to repair, and the absence of brakes. But, in most cases, the worst part of the design is the crude harnessing device. Better harnessing, such as that shown in



Figure 18.6 Head baskets.

Source: World Bank (1983).

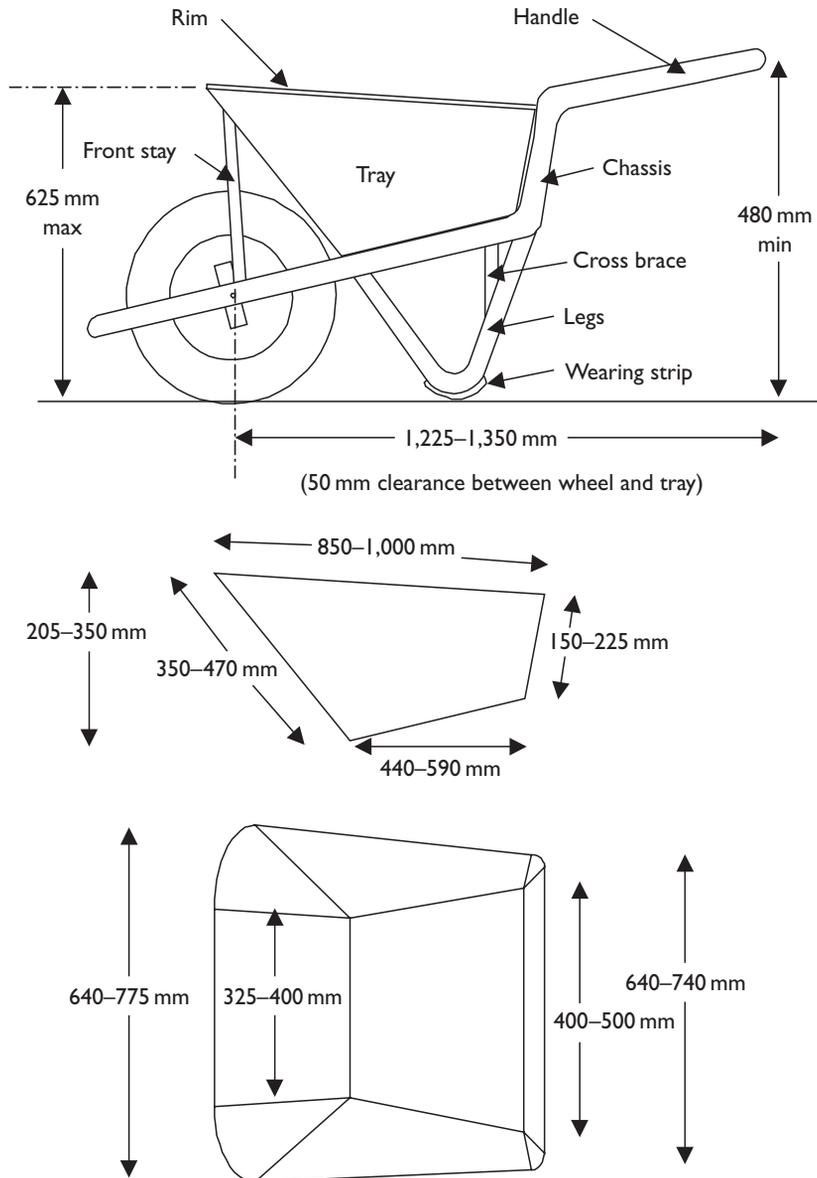


Figure 18.7 Appropriate design of wheelbarrow.

Developed from: ILO (1981).

Figure 18.8, may even often double the output of the animal through reductions in pain and injury, and through more efficient load paths to the cart (Dennis 1996).

Tractor-trailer combinations are very appropriate for labour-based works. One tractor can be used with several trailers. Since the trailer can be unhitched, the expensive tractor does not have to stand idle while the trailers are being loaded, but can be transporting another loaded trailer. However, experience has shown that the trailers need to be very carefully constructed if they are to withstand the rigours of a works site.

Tractor-trailer



Figure 18.8 Improved animal harness.

Source: University of Warwick Development Technology Unit.

Two-wheeled tractor

The two-wheeled, single-axle tractor has played a vital role in the mechanization of agriculture in Asia. It can also be used very effectively for different road works activities when hitched to different specially designed implements, such as trailers, compaction equipment and light scrapers. Advantages are its low cost and robust simplicity, which allows local manufacture and maintenance.

Trucks

Trucks may be appropriate for long haul distances. However, the integration of labour-based and equipment-based methods requires special care. Labourers cannot be expected to excavate and load with the same efficiency as machines. The loading height must be minimized for effectiveness.

Comparison of methods

The technical and economic feasibility of a range of earthmoving methods is illustrated by Figure 18.9.

18.3.6 Organization of work

The organizational structure of labour-based works differs fundamentally from that of equipment-based operations. The size and structure of the organization will vary with the size of the project.

Gang leaders and supervisors

A proven organization structure is to have one gang leader for every 20–25 workers, and one site supervisor in charge of four to five gangs. The gang leader would normally be a worker who has the confidence of his colleagues. He would normally

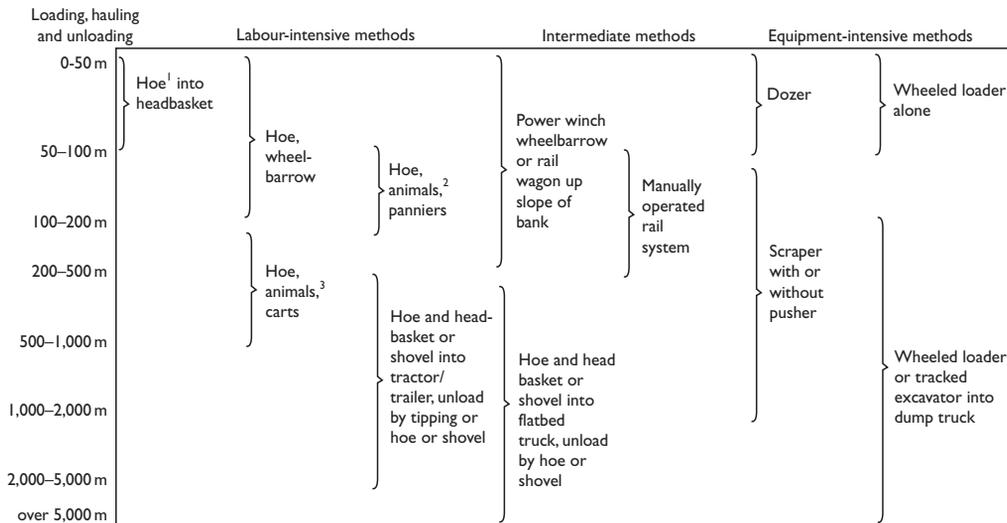


Figure 18.9 Available earthmoving methods.

Source: Edmonds and Howe (1980).

Notes

- 1 The word 'hoe' is used in the study to describe the back-acting tool used for digging in soft soil and for loading, variously named 'powrah', 'mamti', etc.
- 2 i.e. mule, donkey or camel.
- 3 i.e. mule, camel or bullock.

leave the project when his gang terminates its employment, although the best gang leaders could be kept for training to become site supervisors. Site supervisors could be semi-permanent employees kept for the duration of a project, or could be permanent employees moving from one project to another.

For every two to four site supervisors (200–400 workers), a senior supervisor is needed, with good experience and technical knowledge of road works. A technician with engineering training would be suitable for this post. Their job is to solve organizational and technical problems on projects consisting of many small and scattered sites.

Senior supervisors

Finally, there should be a project manager/road engineer to direct two to three senior supervisors. The engineer would be assisted by supporting staff for surveying, accounting, supply and stores, work studies and other administrative matters. The following number of supervisory staff would then be needed for a project employing about 1,000 workers:

Engineer

- one engineer
- three senior supervisors
- 10–12 supervisors
- 40–50 gang leaders.

The recruitment of labour can be done in various ways. It is strongly recommended to involve the local administration and community leadership to the maximum extent

Recruitment

possible. Recruitment should be announced formally and well in advance of the start of the works. If the number of workers applying is greater than needed, a ballot system (selection by lottery) should be applied.

Transport

When casual labourers are employed, transport to and from the site is not normally provided. This restricts the area from which labour can be drawn to 5–15 km from the site. As the work progresses, the distance to walk will become too great for the workers and new ones will have to be employed. The disadvantage with this system is that productivity will decrease because it takes some time for the labourers to acquire efficient work methods. The alternatives are to provide transport for the workers or to arrange labour camps. Both alternatives are costly and bring with them administrative and organizational problems.

Planning and control

For labour-based works, careful planning is particularly important. Planning should be carried out using standard procedures on a monthly, weekly and daily basis. Furthermore, the progress and the use of resources must be controlled. The control process needs to take place at all levels of the organization hierarchy. To enable this, a flow of reports must be established. A good reporting system is essential in a programme with many small sites, since these cannot be visited often by project management.

18.3.7 Payment systems and motivation

Daily payment

Proper motivation is essential to successful labour-based works. Workers paid under task-rate or piece-rate incentive payment systems produce a much higher daily output than daily-paid labour. Daily payment should only be used: (i) when no productivity data for the major operations in infrastructure works is available; (ii) for those activities which cannot be easily measured; and (iii) in countries where incentive payment is forbidden by law or labour agreement.

Task-rate

Task-rate payment is where a fixed sum is paid for a given quantity of work. A daily task should be defined to enable a good worker to finish in five to six full work hours. The worker is then free to go home. The gang leader in charge needs to ensure that the work has been completed satisfactorily, before the worker is allowed to leave the site. Tasks can be given to individuals or groups of workers. Payment by task-rate may be unsatisfactory for activities such as gravelling, where labour and machines have to work together, since the labourer's goal is to finish early. This can result in equipment being underutilized unless it can work independently during the remaining part of the day. Labour and heavy equipment working on the same site on interrelated activities is not recommended, because the machines' much greater output demoralizes labourers. In Ghana, a system has been successfully introduced where labourers who consistently attend work and achieve their tasks over weekly or monthly periods are paid a bonus on top of their task-rate payment.

Piecework

Piecework is where workers are paid a fixed sum per unit of output, for example, one US dollar per cubic metre of hard soil excavated. The worker decides how much to produce and earn. The payment per unit of output has to be determined very carefully and should only be introduced after enough reliable productivity data have been collected. The advantages of this system are that the unit costs are lower than the unit costs achieved with task work, and that productivity is high. The system works well with migrant labourers who lack the incentive to finish early because they live

in a labour camp. Government agencies often have administrative difficulties in paying on a piece work basis, but contractors are more flexible and able to use this method effectively. Care has to be taken that the use of piecework does not lead to worker exploitation. If time limits are not set, workers may be tempted to work very long hours to earn more money and over-exert themselves. Negotiations should aim to set piecework rates at such levels that an average worker can earn at least the minimum wage in normal hours.

Output and productivity levels of labour-based projects depend heavily on the degree of confidence of small contractors and/or workers in their employer. It is therefore crucial that labour-based projects have good administrative and financial procedures relevant to working with large numbers of local workers. In particular, there is a need for timely payments supported by good disbursement procedures.

It is imperative to pay wages on time to retain workers' confidence in the employer, whether it be the government or a contractor. A contractor, however, is more likely to be suspected of mismanagement of funds, and work will be seriously disrupted if payment is not forthcoming at the right time.

Disbursement procedures are necessary to ensure the availability of funds at the right time and at the right place. Small labour-based contractors should receive regular payment to pay their casual labourers. A system of monthly advance payments, deductible from payments certificate, is particularly helpful when new small contractors are becoming established. Subsequently, a system may be introduced where work certificates serve as bank guarantees for overdrafts.

Administrative
and financial
procedures

Timely
payment

Disbursement
procedures

18.3.8 Training

Training needs are most crucial at the level of senior supervisors and site supervisors. Senior supervisors should already have the skills necessary for inspecting workmanship, planning and programming, and assuring reporting accuracy. Field supervisors should be fully competent in the work methods to be used, as they will need to give daily instructions on site, and manage the gang leaders and the workers. Contractors' supervisors need to receive similar training as field supervisors employed by government. Training courses may also be needed for storekeepers, drivers and mechanics, and special courses may be necessary for maintenance inspectors. Gang leaders and workers would normally be trained on-the-job, rather than through formalized training. This should take the form of both instruction and practical demonstrations. However, a number of projects have also started gang leader training in a more formal and structured manner. Small local contractors normally need to be trained in cost accounting, estimating and bidding. In addition, they may need training in contract documentation, including rights and obligations of the contractor, specifications, drawings, and payment procedures. Staff of the road administration and local consultants are likely to require training in contract management and supervision.

Trainees

The requirements of labour-based road works programmes means that it is necessary for training courses to be specifically adapted. Such courses should have a large practical component and emphasize worker management and organizational aspects of the work. For the higher managerial levels, relevant international courses have been developed for engineers, senior technicians, contractors and consultants.

Courses

The ILO also regularly organizes fellowships for managers and decision-makers to visit successful labour-based programmes in Africa and Asia. Training materials on this subject are also being integrated into the course materials of civil engineering faculties of some universities. Courses can be based on general written and audio-visual material developed by ILO (ILO 1991). If trained staff do not immediately have the possibility to apply their knowledge and skills, they become demotivated and training resources are wasted.

18.4 Equipment management

18.4.1 Basic principles

Principles of management

The costs of owning and operating equipment are high, and often represent the largest single component of the cost of the works. Lack of working equipment has been cited as one of the most significant factors in road works operations that are inefficient (World Bank 1988). It is therefore crucial that equipment operations are well managed. Equipment management is a specialized operation, and should therefore be supervised by mechanical engineers. One basic principle is that equipment should normally be managed as an autonomous unit, separated functionally from the works unit. The equipment unit should be responsible for the procurement, repair and disposal of all vehicles and equipment, and should rent these to the works units on commercial terms. Detailed guidance on the setting up and operating such an equipment unit has been given by the World Bank (Lantran and Lebussey 1991), and this section provides a summary of their recommendations.

Commercial management

The commercial management of equipment requires that rental rates are charged for the use of equipment, and that these are sufficient to cover all of the operating costs of the equipment units. Income from rental provides a fund from which replacement and additional equipment can be procured in the future. Such an approach tends to encourage optimum use of equipment because of the following reasons:

- The costs of operating equipment are explicit (and large), so there is pressure on works units to utilize only equipment that is actually needed at any time, and to obtain the maximum productivity from rented equipment, otherwise unnecessary expenditure is incurred.
- No rental is received by the equipment unit for equipment not being used by works units, so there is an incentive to repair broken down equipment quickly and get it back into operation, and also to dispose of old and unwanted equipment.
- The equipment unit can develop a long-term strategy for equipment supply, and can respond to demands from customers in a flexible manner.

18.4.2 Organizational arrangements

Organizational independence

Even for public sector organizations, the equipment unit needs to be independent of government procurement rules and should, for example, be able to operate its own bank account. This enables spare parts to be procured in an efficient manner. Speed

of delivery is important to avoid keeping contractors and works units waiting because of slow repair times, and to avoid rental revenue being lost. The equipment unit needs to be able to negotiate and set up arrangements, with suppliers, that are economical and efficient. Similarly, the unit must have the ability to set rental rates that reflect the true cost of owning and renting equipment. It also requires the flexibility to adjust these to reflect changes in the financial environment, such as inflation and exchange rate fluctuations, if adequate funds are to be generated for equipment replacement. The unit should also have the ability to dispose of uneconomic equipment easily, and to retain any income generated from this.

Equipment maintenance operations can be conveniently organized according to their technical complexity into a number of 'levels' that are suitable for undertaking at different locations. A typical example is shown in Box 18.1. The actual size, number and role of the workshops will depend on the size of the road administration and the types of works for which it is responsible.

The size of the equipment fleet should depend on the current and likely future demand. It will also depend on local circumstances and transport conditions. The organization structure in Box 18.1 can accommodate 125–250 machines and

Organization
of mainten-
ance and
repair

Fleet size

Box 18.1 Organization of equipment maintenance operations

Level 1

Maintenance carried out at the road works site, and at local service facilities if distances to work sites are short – control of lubricants, cooling fluids, pressures in tyres, turbochargers, etc., and visual inspection of machines.

Level 2

Work sites and local service facilities, as above – replacement of oil filters, hoses and accessories that do not require tuning or specialized tools (e.g. alternators, starters, water pumps), and minor repairs.

Level 3

Fixed workshops, including service facilities or regional workshops with towing and maintenance equipment – replacement of major components, such as engines and transmissions, and of accessories requiring special tools.

Level 4

Large fixed workshops (central or regional) – repair of major components not requiring major milling or manufacturing, such as replacement of pistons and cases, and tuning of injector pumps.

Level 5

Central workshops with highly qualified mechanics, or authorized dealers – repair of major components, with the ability to undertake major refurbishment operations, such as straightening cylinder heads and reboring.

Adapted from: Lantran and Lebussy (1991).

250–450 vehicles. In some situations, it may be preferable to set up several regional equipment units instead of a centralized operation. However, because of the fixed costs associated with workshops and offices, there is a minimum size of operation that is viable from a business point of view. This minimum number is likely to vary from over 100 machines and vehicles in large areas where equipment utilization rates are moderate, down to about 60 machines in highly competitive environments. There is also considerable scope for contracting out workshop activities, particularly for Level 4 and 5 operations.

Spare parts management

Spare parts management is also a crucial element of equipment unit operation. The policy should normally be to maintain an adequate stock of fast moving parts, and to arrange expeditious supply of other parts from the manufacturer or from dealers.

18.4.3 Costing and pricing

Rental rates

Setting rental rates is a key issue for an equipment unit, and the principles to be considered when doing this include

- true costs should be known;
- revenues should cover costs;
- revenues should be easy to calculate;
- rental rates should encourage customers to return machines that they are not using; this will allow the unit to allocate these to other customers.

Cost control

Cost control is essential for the equipment unit since, if independent, it would not be able to rely on subsidies from its parent road administration. Separating the equipment unit functionally from the road works organization facilitates this, since costs are explicit and not mixed up with the costs of other operations. The equipment unit will need to have a system for collecting cost data, and for classifying these costs by the type of machine to determine appropriate rental rates. There are three main categories of cost: running cost, depreciation and non-allocable general expenses. Running costs are relatively easy to monitor and control, but the other categories are discussed further.

Depreciation

Financial depreciation involves collecting enough revenue during the life of a machine so that it can be replaced when needed. As such, it is recommended that the depreciation account is separated from the current management account, and that revenue is recorded in the currency used for the purchase of the machine. This enables exchange rate variations to be reflected. Yearly depreciation forecasts should be consistent with realistic life-spans and the replacement cost of new machines, and these should feed back into the rental rates. Realistic residual values should also be taken into account, and revenue from equipment disposal should be placed into the depreciation account. Such an approach also makes it easier to keep track of costs and revenues, and to update rental rates. It also provides a pool of money for equipment replacement.

General expenses

It is not necessary to allocate all costs directly to particular machines. General expenses and labour costs should be identified, and allocated to machine types using simple formulae. In a rental rate calculation, the cost of mechanics and lubricants can be related to the replacement cost of an equivalent new machine.

18.4.4 Contractor development

An approach to developing domestic equipment-holding contractors is described by Lantran (1990). A key aspect of this is the provision of supportive measures for the acquisition of equipment. Many new contractors start as labour contractors, who generally possess little or no equipment. The development approach should normally assist contractors in obtaining loans for the acquisition of equipment. Items of light equipment, which are indispensable for labour-based road works, normally include a lorry or tractor-trailer combination for hauling over longer distances, a pick-up for supervision and transport of light items, and compaction equipment. Assistance will also be needed for the development of in-house expertise for equipment management. Training for equipment management should comprise costing practices, preventive maintenance and on-site repairs. An investigation should be made of the manufacturing and repair capability, as well as the extent and flexibility of the second-hand market in the country concerned. For repair and maintenance services, deals should be concluded with equipment suppliers/agents or plant pools. Such arrangements can often be made by a contractors' association.

References

- Antoniou, J., Guthrie, P. and de Veen, J.J. (1990) *Building Roads by Hand: An Introduction to Labour-based Road Works*, Burnt Hill: Longman for the International Labour Office.
- Bentall, P., Beusch, A. and de Veen, J.J. (1999) *Employment-intensive Infrastructure Programmes: Capacity Building for Contracting in the Works Sector*, Geneva: International Labour Organization.
- Beusch, A. and de Veen, J.J. (1991) *International Course for Engineers and Managers of Labour-based Road Works and Maintenance Programmes*, Volumes I and II, Geneva: International Labour Organization.
- Dennis, R.A. (1996) *Guidelines for Design, Production and Testing of Animal-drawn Carts*. London: Intermediate Technology Publications.
- Edmonds, G.A. and Howe, J.D.F.G. (1980) *Roads and Resources: Appropriate Technology in Road Works in Developing Countries*. London: Intermediate Technology Publications.
- Edmonds, G.A. and de Veen, J.J. (1991) *Technology Choice for the Works and Maintenance of Roads in Developing Countries: Developments and Guidelines*, Geneva: International Labour Organization.
- ILO (1981) *Guide to Tools and Equipment for Labour-based Road Works*, Geneva: International Labour Organization.
- ILO (1991) *Training Videos for Labour-based Road Construction and Maintenance, and Video Scripts (set of 16 Training Modules on 12 Videos, Complete with Scripts, VHS/PAL format)*, Geneva: International Labour Organization.
- ILO (2000) *Employment-intensive Investment in Infrastructure: Jobs to Build Society*, revised edition, Geneva: International Labour Organization.
- Jones, T.E. and Petts, R.C. (1991) 'Maintenance of minor roads using the lengthman contractor system', in TRB, *Fifth International Conference on Low Volume Roads, Raleigh North Carolina, 19–23 May 1991. Transportation Research Record 1291*, Volume 1, Washington, DC: Transportation Research Board, National Research Council, 41–52.
- Karlsson, L.S. and de Veen, J.J. (1992) *Guide to the Training of Supervisors for Labour-based Road Works and Maintenance*, Geneva: International Labour Organization.
- Lantran, J.M. (1990) 'Developing domestic contractors for road maintenance in Africa', *Sub-Saharan Africa Transport Policy Program, Road Maintenance Initiative, Contracting*

- Out Road Maintenance Activities*, Volume I, Second edition, Washington, DC: The World Bank.
- Lantran, J.M. and Lebussey, R. (1991) 'Setting up a plant pool', *Sub-Saharan Africa Transport Policy Program, Road Maintenance Initiative, Contracting Out Road Maintenance Activities*, Volume III, Washington, DC: The World Bank.
- Stock, E. and de Veen, J.J. (1996) 'Expanding labor-based methods for road works in Africa', *SSATP Working Paper No. 22*, Washington, DC: The World Bank.
- Tajgman, D. and de Veen, J.J. (1998) *Employment-intensive Infrastructure Programmes: Labour Policies and Practices*, Geneva: International Labour Organization.
- de Veen, J.J. (1980) *The Rural Access Road Programme: Appropriate Technology in Kenya*, Geneva: International Labour Organization.
- de Veen, J.J. (1991) *Organizational Requirements of Efficient Rural Infrastructure Programmes*, Geneva: International Labour Organization.
- World Bank (1983) *Labour-based Works Programs*, New York: Oxford University Press.
- World Bank (1988) 'Road deterioration in developing countries: causes and remedies', *World Bank Sector Policy Study*, Washington, DC: The World Bank.

Maintenance



Towed grader. (Photo: Bent Thagesen)

Maintenance management

Richard Robinson

19.1 The road network as an asset

Roads in all countries represent an important national asset. For example, the asset value of the road network in the emerging countries of Central and Eastern Europe is estimated to be in excess of US\$550 billion. For many countries, the asset value will be comparable with that of some of the world's largest companies. The value of the asset managed by the Highways Agency in the United Kingdom is comparable with that of IBM; the Japan Road Public Corporation manages assets roughly equal to those of General Motors (Heggie and Vickers 1998). Management of a road network asset of such a value requires adoption of the most careful management practices commensurate with those adopted by the most successful businesses.

Asset value

Despite the importance of managing road networks in a business-like manner, many countries find this to be a difficult challenge. A number of reasons are responsible for this, including

The maintenance problem

- insufficient funds, including a lack of foreign exchange – caused by the weak macroeconomic situation in many countries, but compounded because available funds are often diverted to politically visible capital projects, or utilized poorly, sometimes to support large unproductive labour forces;
- shortage of qualified staff – even in those countries with sufficient engineers, their skill base tends to be in the areas of design and construction, rather than road operation and maintenance;
- absence of machines and spare parts – this is partly related to the lack of funds, but also to poor equipment management practices and bureaucratic procedures for procurement of both new equipment and spares;
- deficient institutional arrangements and management capability – many road administrations are poorly focused, with too many responsibilities; emphasis is often put on works execution, while management is often neglected.

19.2 Road network management

19.2.1 Aims

Road asset management has the purpose of maintaining and improving the existing road network to enable its continued use by traffic in an efficient and safe manner. In addition, appropriate management must also take into account issues of effectiveness, and

Purpose, activities and impacts

concern for the environment. Road management can be seen as a process that is attempting to optimize the overall performance of the road network over time (Robinson *et al.* 1998). In other words, the process may be seen to comprise a number of *activities* (or measures) that will have *impacts* (or effects) on the road network. Impacts include those on the following:

- level of service or road conditions;
- national development and socio-economic issues;
- road user costs;
- accident levels and costs;
- environmental degradation;
- road administration costs.

19.2.2 Activities

Asset management classification

It is convenient, for management purposes, to group activities under the headings of ‘operations’, ‘maintenance’, ‘renewal’ and ‘development’ (National Asset Management Steering Group 2000). These headings can be used for all types of road assets, including carriageways and shoulders, bridges and structures, footways and cycle tracks, street lighting, road signs and furniture. Activities are described in more detail in Chapter 20. The present chapter focuses on the management of the following for carriageways.

- Maintenance
 - routine (cyclic, reactive, winter and emergency)
 - periodic (preventive, resurfacing)
- Renewal (overlay, pavement reconstruction).

Management of maintenance and renewal

Maintenance management of road networks can be considered to have the following objectives (Hooper 2001):

- Network safety
 - complying with statutory obligations;
 - meeting users’ needs.
- Network serviceability
 - ensuring availability;
 - achieving integrity;
 - maintaining reliability;
 - enhancing quality.
- Network sustainability
 - minimizing cost over time;
 - maximizing value to the community;
 - maximizing environmental contribution.

19.2.3 Management functions

The functions involved in maintenance management can be described under four headings (Robinson *et al.* 1998) Functions

- *Strategic planning* – long-term decisions affecting the whole of the road network, undertaken primarily for the benefit of senior managers and policy-makers.
- *Programming* – determining those parts of the road network where work can be undertaken with available resources in the next budget period.
- *Preparation* – design of works for individual sections of road, issuing of contracts or works orders for works that have a budget commitment.
- *Operations management* – managing and supervising on-going works in individual subsections of road.

In simple terms, maintenance management aims to get the *right* resources (people, materials and equipment), to the *right* place on the road network, to carry out the *right* maintenance or renewal work, at the *right* time. In many ways, maintenance management presents a greater challenge to the road engineer than works execution. Most maintenance works are relatively straightforward and the skills easily acquired through well-organized training. Maintenance engineers need a detailed knowledge of all maintenance techniques, but supervision of the majority of maintenance works ('operations management') should be delegated to a foreman or technician. The time and skills of the maintenance engineer are better utilized by being concerned with planning and programming (including budgeting), and the monitoring of maintenance operations. The remainder of this chapter focuses on the management function of 'programming'.

Role of the
maintenance
engineer

19.2.4 Management cycle

All maintenance management activities can be carried out by following the steps of the 'management cycle', which is a concept used in business management. The management cycle for 'programming' is illustrated in Figure 19.1. In this case, the aim is to prepare a work programme that can be carried out next year for the budget that is likely to be available. Note that 'implementation' involves the submission and approval of the programme, not the carrying out of the works themselves. The example indicates that each step in the cycle requires access to information to enable the step to be undertaken. Information is central to the management cycle.

Steps in the
cycle

Figure 19.1 also provides the framework used in this chapter for discussing maintenance management. In this case, the *aim* is to 'determine the work programme that can be carried out with next year's budget, that is, 'programming'. The remainder of the steps in the management cycle are discussed under the following headings:

Framework
for discussing
maintenance
management

- Network information
- Assessing needs
- Determining options
- Choosing actions
- Implementation
- Monitoring and audit.

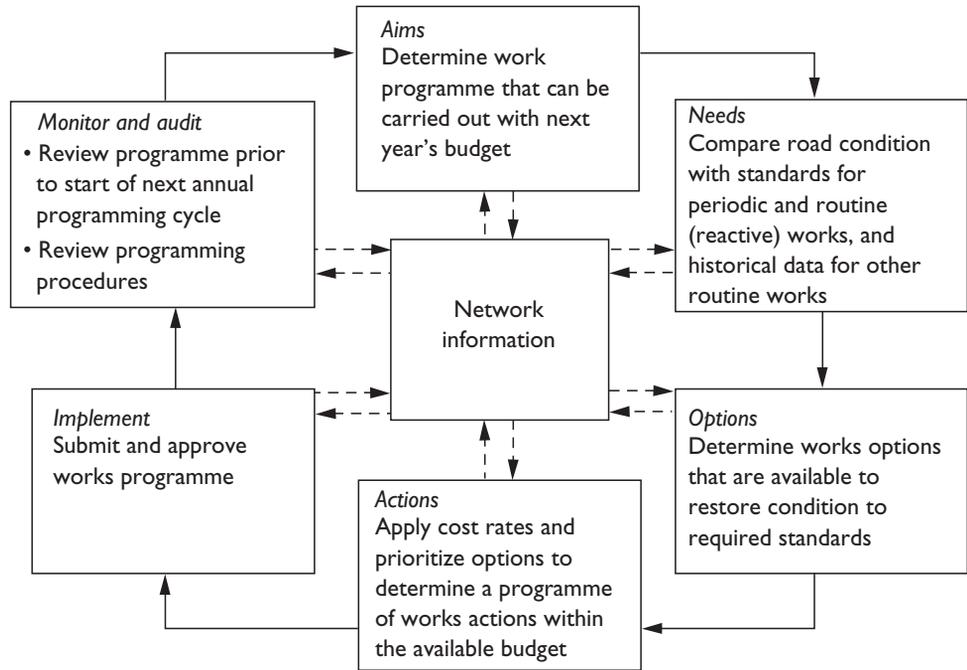


Figure 19.1 Example of the management cycle for programming.

More details about road maintenance management are provided in *TRL Overseas Road Note 1* (Robinson 2003). Additional considerations for very low-volume roads are provided in *TRL Overseas Road Note 20* (TRL and DFID 2003).

This chapter also briefly discusses information systems, since it is now common practice for computer-based systems to be used to assist with the programming function.

19.3 Network information

19.3.1 Classification

Information about the asset being managed is crucial for road maintenance management. Basic information about the length and characteristics of the roads to be managed is obviously essential. However, information on the nature of roads and traffic is also needed because the levels of service and maintenance standards that are appropriate will vary depending on these. Roads are, therefore, normally allocated to categories, which form a hierarchy. A common set of hierarchies is as follows:

- *Arterial* – the main routes connecting national and international centres, with relatively high levels of traffic, speeds and average trip length.
- *Collector* – roads linking rural areas to adjacent urban centres or to the arterial network, with traffic flows and trip lengths of an intermediate level.
- *Access* – the lowest level of road in the network, with low vehicle flows and short trip lengths, and with substantial proportions of total movements likely to be by non-motorized traffic and pedestrians.

19.3.2 Inventory

Information relating to classification and standards must be related to the geographic location of any item on the road network. Thus, a system of network referencing is fundamental to road management. The network is usually broken down into a series of links and sections, each defined by a unique label. The list of links and sections provides a roads register, or gazetteer, which defines the entire road network. It is convenient if the start and finish of sections are identified physically on the road, and marker posts, often at kilometre intervals, are often used for this.

Network
referencing

An item inventory is also needed which lists, against each section, details of the physical characteristics of the road. The level of detail recorded in the inventory may depend on the importance of the road in the hierarchy. Generally, the inventory should be as simple as possible and not overloaded with unnecessary information. An inventory typically contains information such as the following:

Item inventory

- *Sections* – the length of each section in the network.
- *Cross-section* – the width of the carriageway and shoulders, with information on whether side ditches are present.
- *Pavement* – the type, thickness and, if possible, the age of the pavement on the carriageway and on the shoulders.
- *Alignment* – the chainage of characteristic points in the alignment, such as the location of crossroads, culverts, bridges and sharp curves; details of steep gradients and the radii of sharp curves may also be recorded.
- *Structures* – the type and dimensions of major bridges, culverts and retaining walls.
- *Furniture* – information on road signs, guard rails, lighting and other features.
- *Soils* – information about the soil type along the road (clay, sand, rock, etc.), and location of identified deposits of road materials.
- *Land use* – such as town, village, woods, farmland.

Inventory data may often be collected by driving slowly along the road and stopping for measurement of characteristic cross-sections. Chainages can be recorded on the car's trip meter. Information on pavement and structures can be obtained by inspection. However, some testing may be necessary, depending on the application to which the inventory will be put. Horizontal curvature can be determined with a compass, or in relation to the turning angle of the car's steering wheel. Gradients can be measured by means of a simple fall meter. The collected information can be recorded in tables or on schematic road plans, but most inventory systems for road maintenance are computer based.

Recording

Although the preparation of the inventory is a once-only activity, it is very important that it is kept up-to-date. Information on changes to the network, such as new surfacings and reconstruction works, need to be entered into the register, otherwise its usefulness is reduced and assessing maintenance needs is made more difficult.

Updating

19.3.3 Traffic

An important aspect that affects the asset value will be the amount and type of traffic using the road, and this needs to be measured. The need for some maintenance

Need for
traffic data

activities will depend on the traffic volume and the distribution of axle loads on sections of the road network. Traffic and axle loads may be used to define different road hierarchies in the classification, between which maintenance standards and thresholds will vary.

Traffic volume On roads with light traffic, it is normally sufficient to make a rough estimate of traffic using ‘moving observer counts’. These involve measuring traffic whilst driving along the road. Where traffic is heavier, counts are made every few years at selected count stations. Traffic in intervening periods and future traffic are estimated by use of growth factors. Traffic counting and forecasting are discussed in Chapters 3 and 6. High accuracy of traffic measurement is seldom necessary for maintenance management purposes.

Axle loads The distribution of axle loads is needed to design some periodic maintenance treatments and renewals. Measurements are normally made using portable weighing scales. Measurements are normally made only in association with specific planned maintenance actions. If the traffic volume is low, all passing vans, trucks and buses are weighed on an ordinary week day. If the traffic volume is large, then a sample is normally taken. For example, every tenth passing commercial vehicle is stopped and weighed. Under no circumstances should the survey be limited to vehicles with large, visible loads. The results of the axle load survey are transformed into cumulative equivalent standard axles (ESA) for each category of commercial vehicle (see Chapter 3).

19.4 Assessing needs

19.4.1 Defects

Types of defect The road register and item inventory contain data providing basic information about the road network to be managed. These data do not change, or change very slowly, over time. To assess maintenance needs, information is needed about where the network is defective. Road condition does change, sometimes rapidly, over time. The comparison between measured road condition and pre-defined standards, or intervention levels, provides a basic statement of shortfall in serviceability, which can be translated into maintenance need. It is convenient to characterize defects under the following headings:

- Paved roads
 - roughness (unevenness);
 - surface distress (rutting, cracking, spalling, pot-holes, etc.);
 - structural adequacy;
 - pavement texture and friction.
- Unpaved roads
 - roughness, including corrugations;
 - surface distress (loss of camber, rutting, pot-holes, ravelling or loose material);
 - gravel loss;
 - dust.

Defects can be assessed using manual or mechanized methods. Visual inspections are normally used to record all conspicuous defects of pavement, shoulders, ditches, culverts, slopes and road furniture. Mechanized data are more repeatable, reproducible and, generally, can be collected more efficiently. Mechanized data collection provides scientifically based techniques to assist in the determination of causal factors of defects, to monitor changes in condition, to assess strength and to help determine appropriate treatments. However, mechanized measurements are usually limited to the pavement. Detailed descriptions of defects and the methods for assessing them can be found in standard texts, such as Robinson *et al.* (1998), Haas *et al.* (1994) and Shahin (1994).

Defect
assessment

19.4.2 Strategy for data collection

A variety of strategies can be used for collecting data on a regular basis, and different approaches will be appropriate for road networks of different characteristics managed by different road administrations. Examples of strategies are shown in Box 19.1. Ideally, a cost–benefit analysis should be undertaken to determine the survey strategy that is most appropriate, and the optimal level of data detail to be collected. However, whereas it is relatively straightforward to determine the cost of the different strategies, it is not easy to estimate the benefits to be derived from having information at the resulting different levels of detail.

19.4.3 Visual inspections

In many situations, a visual condition survey of at least part of the road network will be carried out at least once a year to assist in determining maintenance needs for the next budget period. The most appropriate time of the year for this will depend on the

Timing and
responsibility

Box 19.1 Typical strategies for collecting condition data

Strategy A

Summary condition data are collected across the whole network each year. These data are used for strategic planning and programming purposes. The programming exercise then collects more detailed data on those sections where works are likely to be undertaken. Very detailed data are then collected on some of the sections for which designs are produced, or for which works are undertaken.

Strategy B

Relatively detailed data are collected across part of the network on a rolling programme, perhaps with a cycle of three to five years. Each year, programming decisions are taken either using current data for individual sections, if available, or by projecting forward condition data from previous years to give an estimate of current condition across the whole network.

Other strategies are also possible (see, e.g. Robinson *et al.* 1998).

climatic conditions. The drainage system should ideally be inspected in wet conditions, since this can only be evaluated satisfactorily when there is water present. In climates with freezing conditions, inspections should be carried out immediately after the winter season, when thawing conditions can result in significant damage to the road infrastructure. It is useful if the engineer responsible for road maintenance participates personally in at least some of the visual inspections. This will ensure that maintenance works can be planned effectively, based on personal familiarity with the road network, and that the quality of work can be monitored.

Recording

Normally when inspecting a road section, the road is divided into subsections, typically 100 or 200 m in length. The marker posts related to the road register can be used as a reference, if they are present. For each distress mode, the extent and severity of the defect are recorded. Recording of defects should be supplemented by an assessment of their possible causes. Knowledge of causal relations is fundamental for the selection of appropriate repair methods. Figure 19.2 shows an example of a form used for the inspection of paved roads, and Figure 19.3 shows equipment that can be used for measuring deformation, such as rutting.

19.4.4 Mechanized measurements

Cost-effectiveness

Many of the mechanized methods for defect assessment are relatively sophisticated and expensive, and the benefits of their use should exceed the cost of their procurement. Some simple devices are described here.

Roughness

The most important functional defect of a road is the roughness of the riding surface. The 'bump integrator' (BI) is a simple device for measuring roughness and it is widely used throughout the developing world. BIs use a response-type approach to measure the sum of relative displacements between axle and body of a running vehicle. Instruments can be installed in an ordinary passenger car. Roughness is measured with a vehicle-mounted BI. An outline diagram for this equipment is shown in Figure 19.4. It consists essentially of an integrator unit attached to the vehicle's floor and connected to the rear axle. The vehicle's rear axle is connected by a cord to the cylindrical drum on the integrator unit, tension in the cord being maintained by a return spring inside the drum. The drum is attached to a shaft operating a six-lobed cam through a one-way clutch. The cam closes a pair of contacts once for each inch of integrated downward movement of the wheel, the accumulated movement being recorded on an electromagnetic or electronic counters, which are fitted into a panel board mounted in the vehicle. A changeover switch on the panel board enables measurements to be recorded alternately on different counters. The power supply is taken from the vehicle's 12-V battery. The distance travelled along the road is measured by an accurate odometer ('rally meter') fitted to the panel board of the vehicle. An accelerometer can also be used as a simple response-type roughness meter. The accelerometer is installed in a passenger car simply by fastening the instrument to the body of the vehicle. The accelerometer measures the vertical acceleration of the body at very frequent time intervals. The results may be recorded in different ways.

Roughness calibration

The results from response-type instruments are related not only to the longitudinal profile in the wheel path, but are also dependent on the characteristics of the measuring vehicle. Calibration of the instrument and the vehicle are necessary if the results are to be utilized for assessing maintenance needs. Calibration also enables

Surface condition survey	Road no.		Form no.	Date		Inspector
	Start point	Road width		Surfacing	Shoulder	
Start chainage						
Direction towards						
Cracking						
• Type						
• Intensity						
• Position						
• Width						
• Extent						
Rut depth						
Agg. polishing						
Surface texture						
Pot-holes/patching						
Notes:						
Direction towards						
Cracking						
• Type						
• Intensity						
• Position						
• Width						
• Extent						
Rut depth						
Agg. polishing						
Surface texture						
Pot-holes/patching						
Notes:						

Figure 19.2 Inspection form for paved roads.
Adapted from: TRL and DFID (1999).

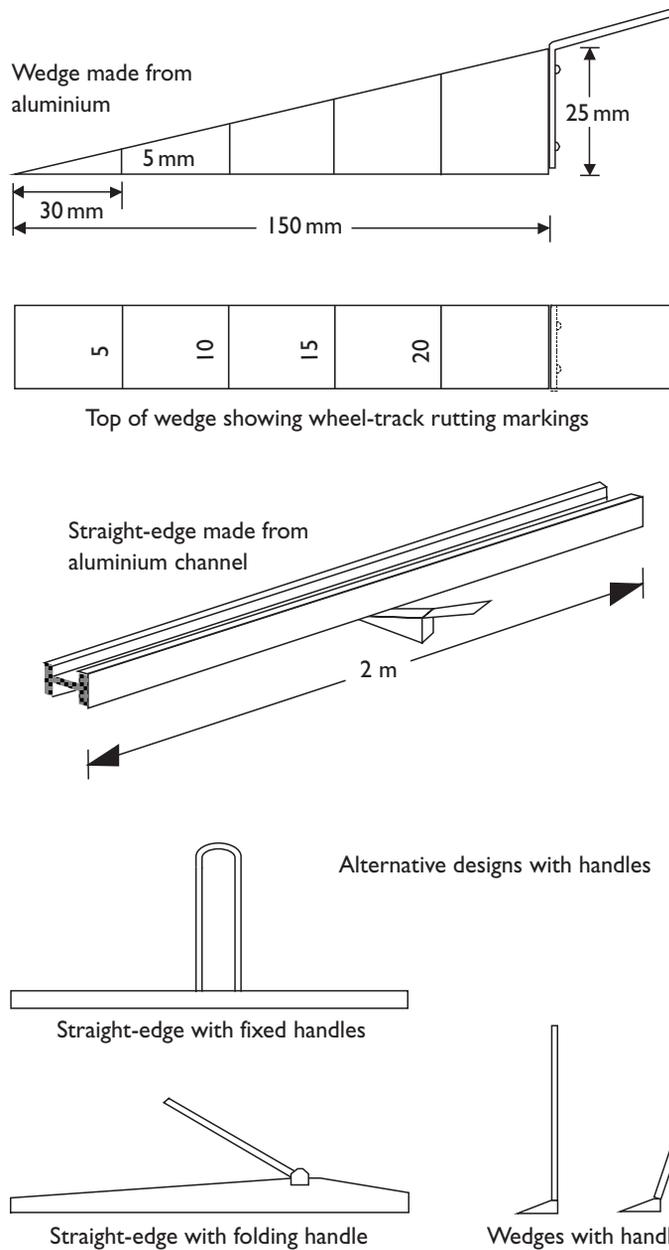


Figure 19.3 Deformation gauge.

values obtained from a response-type instrument to be converted to units of the International Roughness Index, which is the standard measure of road roughness (TRL and DFID 1999).

Deflection

The structural condition (bearing capacity) of a road pavement is often evaluated based on the road surface deflection when loaded with a heavy standard wheel load. The ‘benkelman beam’ (TRL and DFID 1999) is a simple hand-operated device for measuring pavement deflection, and it is used in many countries as a substitute for more sophisticated and expensive instruments (see Chapter 15). The ‘falling weight deflectometer’ (FWD) is becoming increasingly popular for evaluating the structural

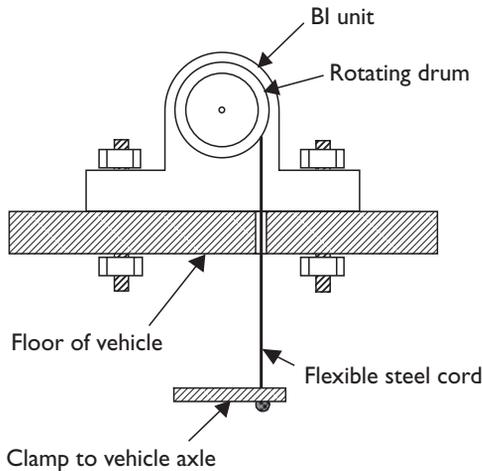


Figure 19.4 BI unit for measuring roughness.

condition of road pavements. Use of the FWD enables the deflection bowl of the pavement under a dynamic load to be measured. The elastic modulus of each pavement layer can then be calculated from the results, provided that the thicknesses of the pavement layers are known.

The DCP (TRL and DFID 1999) can be used for the rapid measurement of the in situ strength of existing pavements constructed with unbound materials (see Chapter 8). DCP tests are particularly useful for identifying the cause of road deterioration when associated with unbound layers; for example, shear failure of the base or sub-base.

Dynamic cone penetrometer

19.5 Determining options

Determining options for action involves selecting possible treatments that can be used to cure defects and to restore conditions to the required level. The use of standard rules for treatment selection ensures that a consistent approach is taken to specifying works throughout the road administration. This helps to ensure that available funds are spent to greatest effect, and that each road and part of the network receives its fair share of the budget. Two fundamentally different types of rules are available:

Treatment selection

- *Scheduled* – fixed amounts of work (such as a quantity in m^2/km) are specified per unit time period (such as one year), or work is specified to be undertaken at fixed intervals of time.
- *Condition-responsive* – work is triggered when condition reaches a critical threshold, known as an ‘intervention level’.

Cyclic routine maintenance works are carried out on a scheduled basis. This approach is used where need is related to environmental conditions, such as cutting back vegetation growth or cleaning culverts. The approach is also appropriate where the deterioration rate is stable over time. Also, where deterioration rates are rapid, such as for the surface of gravel roads, it is impracticable to respond to defects assessed as a result of condition surveys. Many road administrations also

Scheduled treatments

Table 19.1 Example of some intervention levels for paved roads

Defect	Level	Extent (%)	Defect	Extent (%)	Action
Stripping or fretting	Any	<10	—	—	Local sealing
		>20	—	—	Surface dressing
Edge damage	>150mm	>20	—	—	Patch and repair edge
Wheel track rutting	<10mm	—	Wheel track cracking	<5	Seal cracks
				>5	Surface dress
	10–15mm	>10	Non-wheel track cracking	<10	Seal cracks
				>10	Surface dress
	>15	<10	Any cracking	—	Crack sealing
			Cracking only associated with ruts	—	Patch
		Other cracking	—	Patch excess rutting and treat cracks	
	>10	Any cracking	—	Further investigation	

Adapted from: *Overseas Road Note 1* (Robinson 2003).

schedule some periodic works, particularly where the effect of a severe environment outweighs the damaging effect of traffic. Surface dressing is a common example.

Condition-responsive methods

Treatments are triggered, in most condition-responsive methods, whenever one or more defects exceed their respective intervention level. An example of simple rules is shown in Table 19.1. It will be seen that the method requires that the severity and extent of defects is determined. Different sets of rules may be needed for roads classified into different functional hierarchies. This recognizes that the intervention level will depend on the level of serviceability expected from a particular road.

Condition indices

Condition indices combine defects into groups using functional relationships for treatment selection purposes. They provide

- a generic statement about the defectiveness of the four main defect groups of roughness, surface distress, structural adequacy and pavement texture and friction;
- a convenient grouping of defects as an interim step in a calculation or algorithm for determining treatments.

Such an approach is helpful, since most treatments correct one or more of the four defect groups. For example,

- a surface distress condition index might be a function of the defects of surface cracking, fretting and bleeding;
- a structural adequacy condition index might be a function of the defects of structural cracking, rutting and deflection.

Use of condition indices for treatment selection is illustrated graphically in Figure 19.5. However, caution should be exercised when interpreting condition

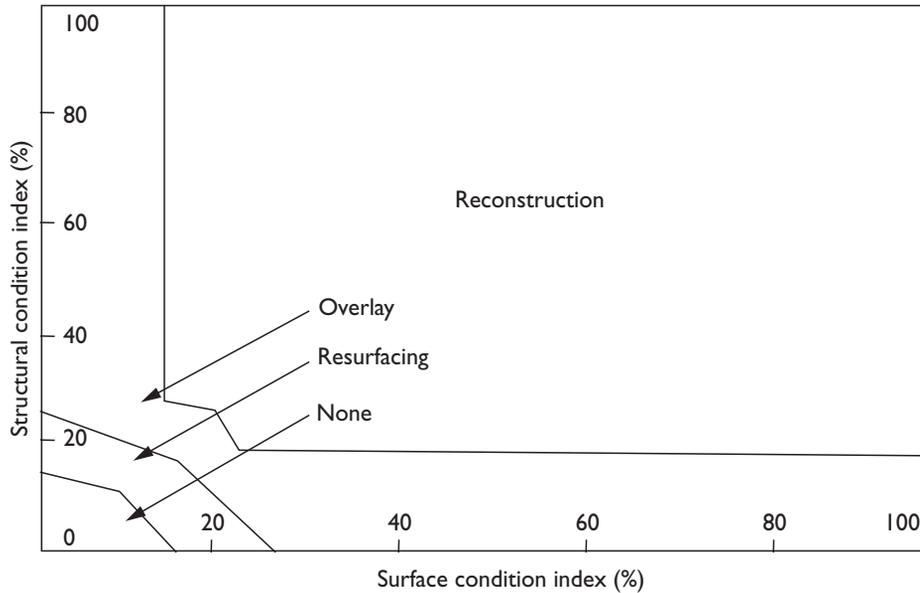


Figure 19.5 Example of use of condition indices for treatment selection.

Adapted from: Phillips (1994).

indices. Condition indices combine surface and structural defects, but these are non-commensurable concepts. Thus, although condition indices are widely used, the lack of a logical basis presents a problem. For example, a bad surface condition could be caused by defects in the surface itself, or by structural defects. Defects in the surface should be treated by surface repair, whereas strengthening is needed to treat structural defects. Thus, the use of condition indices is helpful for making generic statements about pavement condition, for example in a network-level analysis, but they should not be used for detailed treatment design.

Methods of treatment selection are commonly incorporated into computer-based programming systems, commonly known as 'pavement management systems' (PMS). Many systems recommend an appropriate action, and then allow the engineer to interact with the system to take account of local knowledge and modify the treatment choice. A wide variety of treatment selection methods are available in addition to those described here. A fuller description is given by Robinson *et al.* (1998).

Use of
treatment
selection

19.6 Choosing actions

19.6.1 Selection from among options

The treatment selection process may indicate a number of different treatment options for an individual section of road. The options may each have a different cost, and a different treatment life. The process of choosing the appropriate action has to take account of the options available on each section of road, bearing in mind that there is unlikely to be sufficient budget to fund all of the treatments needed across the network. The process therefore involves the two steps of costing and prioritizing.

Table 19.2 Examples of work outputs from activities

Activity	Resource requirements			Materials	Output unit	Range of outputs
	Personnel	Equipment				
Clearing side drains by hand	4-10	Shovels, cutlasses, picks	—	—	m/worker-day	30-60
Clearing side drains by machine	2-3	Grader, shovels	—	—	km/day	4-7
Re-excavating side drains	2-10	Picks, shovels	—	—	m/worker-day	8-15
Clearing culverts	2-4	Shovels, head-pans/wheel barrows	—	—	No./worker-week	2-4
Minor repairs to culverts	2-4	Masons' tools	Cement, aggregate, sand	—	No./worker-week	2-10
Major repairs to culverts	To be assessed for each job				Worker-day	—
Making culvert rings (1 m diameter × 1 m long)	4-10	Moulds, mixer, shovels	Cement, stone, sand, reinforcement	—	No./day	5-10
Grading unpaved surfaces	2	Grader, camber board, spirit level	—	—	Pass-km/day ^a	20-50
Dragging unpaved surfaces	1	Tractor and drag	—	—	Pass-km/day ^a	20-50
Patching bituminous surfacings	5-7	Pedestrian roller or hand rammers, brushes, picks, shovels, watering cans	Premix or gravel, bitumen emulsion, chippings or washed gravel	—	m ³ /worker-day	0.5-0.8

Filling gravel surfaces	5-7	Pedestrian roller or hand rammers, brushes, picks, shovels, watering cans	Gravel	m ³ /worker-day	0.6-1.2
Filling earth surfaces and slopes	4-5	Hand hammers, brushes, picks, shovels	Selected earth	m ³ /worker-day	0.9-1.5
Grass cutting by hand	2-10	Cutlasses	—	m ³ /worker-day	300-800
Grass cutting by machine	1-2	Tractor/mower	—	km/day	10-20
Repairing and replacing traffic signs	2-3	Masons' tools, painters' tools, shovels	Cement, stone, sand, paint, reflective paint	No./worker-day	4-8
Road markings	2-4	—	Road paint	m/worker-day	50-200 (hand painting) 450-500
Stock-piling gravel by hand	10-20	Picks, shovels	—	m ³ /day	300-350
Stock-piling gravel by machine	4	Bulldozer, loader	Gravel	m ³ /day	300-350
Regravelling gravel surfaces	12-20	1 grader, 8 tippers, 1 loader, 1-2 water tankers, 1 roller	Gravel	m ³ /day	300-350
Surface dressing	15-20 ^b	1 distributor, 1 roller, 3 tippers, 1 gritter, 1 loader	Bitumen, chippings	Lane-km/day	2.5-4.0

Adapted from: Overseas Road Note 1 (Robinson 2003).

Notes

a 'Pass-km' is the actual distance the grader travels while working.

b + 10 additional if no loader is available, + 10 additional if no gritter.

19.6.2 Costing

Resource requirements The costs and resource requirements depend on the productivity levels anticipated from different maintenance activities. The productivity of an activity can vary considerably from country to country, but should lie within the limits indicated in Table 19.2. It should be noted that the values in this table make no allowance for time that is not spent actually working, including non-productive time due to broken down or non-available equipment. Non-productive time can build up significantly during maintenance operations, and it should be an aim of management to reduce it to a minimum.

19.6.3 Priority assessment

Need for prioritization Usually, funds allocated to maintenance will be insufficient to meet all the identified maintenance needs. This applies to all countries. An important component of the management cycle for programming is to prioritize maintenance works to determine which can be carried out in the budget period, and which works must be deferred until a later time. A variety of methods are available for prioritization, ranging from simple ranking methods up to complex methods involving optimization.

Ranking methods The simplest method of prioritization is to apply a straightforward ranking to the maintenance works depending on the severity of the defectiveness. A slightly better method is to relate priority to the importance of the different road links and partly to the importance of the different maintenance works, rather than defectiveness. In such an approach, main roads are given higher priority than feeder roads and, because of the importance of drainage, cleaning and reshaping ditches may be given higher priority than resealing pavements. The resulting works are placed in a list in order of their rankings and cumulative costs are applied, starting at the top of the list. Maintenance works with the lowest rankings are rejected until the total costs of the remaining works equals the available budget. This is illustrated in Box 19.2. The main problem with ranking models is that they do not take into account the longer-term consequences of decisions, with the result that constrained budgets are not necessarily spent in the optimum manner.

Consequence analysis Some more sophisticated prioritization methods are summarized in Box 19.3. An important additional benefit of the more sophisticated prioritization methods is that they can be used for consequence analysis. For different budgets, the future condition of the roads, maintenance needs, and road user costs can be evaluated. This information is particularly valuable for decisions relating to programming and strategic planning.

19.7 Implementation

In-house or contract operations Traditionally, implementation of maintenance has been undertaken by the road administration itself. However, increasingly, work is procured under competitive contract arrangements to seek benefits in both effectiveness and efficiency. These issues are discussed in Chapter 22. Whichever method of implementation is chosen, the road administration must still manage and supervise the activities.

Worksheets Implementation involves scheduling works and assigning tasks to the maintenance teams. A schedule is essentially a set of instructions that tell the foremen or technicians supervising an activity how much work is to be done each day, the time it

Box 19.2 Example of a treatment-based prioritization method

The method classifies maintenance treatments according to their impact on long-term deterioration. It also classifies roads according to their importance: those carrying the heaviest traffic loads normally being the most important in the network from an economic standpoint. The method assigns top priority to any strategic roads, since it is considered vital to keep strategic roads in good condition. The remainder of the network is classified by level of traffic on each road, and the traffic hierarchy recommended by the method is shown here.

Road hierarchy	Traffic range (vehicles/day)	Surface type
1	(Strategic roads)	Paved
2	Greater than 1,000	Paved
3	500–1,000	Paved
4	200–500	Paved
5	Greater than 200	Unpaved
6	Less than 200	Paved
7	50–200	Unpaved
8	Less than 50	Unpaved

Priorities are determined using the matrix shown here. Maintenance activities are numbered from 1 (highest priority: emergency maintenance on strategic roads) to 48 (lowest priority: capital works on roads with low levels of traffic). Although a structure for the matrix is recommended in the example, the expectation is that users will define their own priorities to reflect local circumstances within this general framework.

	Road hierarchy							
	1	2	3	4	5	6	7	8
Emergency work	1	7	8	9	10	11	12	13
Cyclic drainage work	2	14	15	16	17	18	19	20
Reactive pavement work	3	21	24	27	30	33	36	39
Periodic preventive work	4	22	25	28	31	34	37	40
Other cyclic/reactive work	5	23	26	29	32	33	34	35
Overlay/reconstruction	6	42	43	44	45	46	47	48

Adapted from: *Overseas Road Note 1* (Robinson 2003).

Box 19.3 Summary of some prioritization methods

General approach

The more sophisticated systems adjust the list of repair works by use of economic and life cycle cost considerations. For each section of road, a range of possible treatments is considered. The approach selects the best combination of treatments across the network as a whole to optimize the use of the available budget. A consequence is that, instead of rejecting expensive maintenance operations, these are replaced by cheaper solutions. For example, resealing may be substituted by patching; overlay may be substituted by resealing. In this way, the maintenance can be stretched to a larger part of the road network and a better total solution is obtained. Increasingly, methods consider road user costs as well as costs to the road administration when making decisions.

Cost-effectiveness methods

Priorities are related to the 'cost-effectiveness' ratio of treatment life to treatment cost. These methods reflect the higher value of a treatment that lasts longer. The methods can be relatively modest in terms of data requirements and computational effort.

Optimization methods

The optimum combination of a number of different works options is selected to achieve a given objective, which might be to minimize the life cycle costs on the network, or to maximize the quality of road condition. Costs normally include both road administration and user costs. These methods provide solutions based on a long-term view of the network, but are demanding in terms of data and computational effort.

Reduced time period analysis methods

These methods address theoretical concerns that life cycle optimization methods base their decisions on long-term predictions, which might turn out to be inaccurate. For example, budgets may change from those predicted, or traffic levels and road conditions may differ significantly from those projected. This type of method addresses these concerns by basing decisions on a limited time period of analysis. Their performance is similar to that of optimization methods, but they are less demanding in terms of data and computational resources.

Adapted from: Robinson *et al.* (1998).

should take and the labour, equipment and materials to be used. Figure 19.6 shows an example of a completed worksheet. The target is the first item entered on the worksheet. The amount of each resource to be used is entered next in the top of each line. The worksheet is issued to the supervisor who enters in the bottom half of each line the progress actually made and the resources used day by day. At the end of the schedule period, the completed form is returned to the maintenance engineer to calculate how much of the production target has been achieved.

WORKSHEET							Activity:		Regravelling		
Section:	23/1079		Sub-section:		34			Period:	12-18 August 2002		
	Unit	M	T	W	T	F	S	S	Total	%	
Target production	km	←		0.5 km	/day	→		3.500	71		
Resources									2.490		
Gravel	m ³	←		350 m ³	/day	→		2.450	65		
		300	300	100	100	400	0	400	1600		
Grader	hr	←		8	/day	→		56	75		
		6	8	8	4	8	b/d	8	42		
Roller (2)	hr	←		16	/day	→		112	70		
		16	16	7	8	16	-	16	79		

Figure 19.6 Example of worksheet.

Adapted from: *Overseas Road Note 1* (Robinson 2003).

19.8 Monitoring and audit

Monitoring serves two main purposes: to check the quantity and the quality of the work being done and to provide data that can be used to improve future maintenance operations. Monitoring involves site visits and desk review.

Site visits are important because they enable the maintenance engineer to become thoroughly familiar with road conditions in the area and to gain first-hand knowledge of the extent and the quality of the maintenance that has actually been carried out. The presence of the maintenance engineer on the spot means that personal advice can be given on problems as they arise, and seeing the engineer regularly on the site should boost the morale of road gangs and improve their standard of work and their output.

Desk review is an office task that involves reviewing all the maintenance documentation; that is, inspection reports, completed worksheets, etc. It provides the opportunity to assess the performance of the maintenance programme and the effectiveness of the management system.

The audit function is, essentially, an integrity check to ensure that work has actually been undertaken and meets the required standards and specifications (technical audit), and that funds have been spent for the purposes actually intended (financial audit). A special unit, which is independent of the main functional branches of the road administration, often carries out both tasks. Auditing is normally done on a sample basis.

Feedback

Site visits

Desk review

Audit

19.9 Information systems

Effective maintenance management requires appropriate information to support management decisions, and the quantitative basis for this is provided through data. The processing of data is facilitated by the use of computer-based systems, and these are increasingly being used routinely in all countries for maintenance management. Computerized systems can be used to assist with the following processes:

Use of computers

- *Network information system* – for the storage and management of inventory data.
- *Planning system* – to support the strategic planning process.
- *Programming system* – to support the programming process (often known as a PMS).
- *Operations system* – to support the operations management process.

Implementation problems

The track record of implementing sustainable systems in many countries is poor. Some reasons for this are listed in Box 19.4. It is recommended that, to overcome some of these problems, a systematic approach is taken to designing and operating systems. Detailed guidance on this is provided in *TRL Overseas Road Note 15* (Robinson 1998).

System procurement

Information systems have been implemented in many developing and emerging countries in recent years. Procurement of a system can be considered as a five-phase process:

- 1 *Commitment* phase (decide to proceed)
 - Obtain commitment from key individuals in the road administration to the system implementation process.
- 2 *Requirements* phase (decide what is wanted)
 - Agree the objective for the system and determine what components the system needs to contain; decisions should be supported by cost-benefit analysis.
- 3 *Specification* phase (decide what is needed)
 - Identify users of the system and the outputs that they will require to support them in their management decision-making.
 - Identify data needs and models required to produce these outputs.
- 4 *Procurement* phase (choose the best solution)
 - Identify appropriate software, together with hardware and operating system requirements necessary to support it.
- 5 *Operations* phase (make the system work)
 - Implement the chosen system
 - Initial and on-going training
 - Managing operation of the system.

Box 19.4 Problems with system implementation*Socio-political factors*

- Client attitudes
 - lack of genuine commitment to implementation project;
 - expectation of high-tech solutions when, in fact, simple solutions are more appropriate and sustainable.
- Cultural issues
 - problems of introducing modern management practices into cultures that are gerontocracies, or where inter-ethnic problems exist, or where nepotism and favouritism are prevalent;
 - traditional behaviours where excuses must be found to avoid blaming individuals.
- Economic and financial problems
 - weak local economies and foreign exchange shortages preventing the purchase of even basic commodities needed to support the system;
 - local budgets dominated by the payment of staff salaries, with residual funds being insufficient to pay for maintenance to be carried out.

Institutional issues

- difficulty with consultants staffing teams with appropriately qualified and experienced individuals;
- local and counterpart staff positions not filled, or filled by staff with insufficient experience;
- training
 - operational requirements preventing local staff being released for training;
 - over-ambitious training programmes with instructors being inadequately prepared;
 - insufficient follow-up training and revision.

Technical factors

- Inappropriate and unrealistic terms of reference for system implementation.
- Focus of attention on the procurement of new equipment rather than on the management and systems needed for its maintenance and repair.
- Deficient computer facilities and inadequate availability of hardware.
- Poor availability of existing data.
- Systems being too complicated and demanding to be sustainable with local resources.

Adapted from: Robinson and May (1997).

Key issues Past experience suggests that the following are key issues that must be met when implementing systems to support the maintenance management process:

- The system must be adapted to suit the local conditions.
- Existing methods and procedures that cannot be used unchanged should be modified rather than replaced with new ones.
- Sufficient, well trained and motivated staff must be provided.
- The introduction of the system must be properly supervised.

References

- Haas, R., Hudson, W.R. and Zaniewski, J. (1994) *Modern Pavement Management*, Malabar Fa: Krieger.
- Heggie, I.G. and Vickers, P. (1998) 'Commercial management and financing of roads', *World Bank Technical Paper No. 409*, Washington, DC: The World Bank.
- Hooper, W. (ed.) (2001) *Delivering Best Value in Highway Maintenance: Code of Practice for Maintenance Management*, London: Institution of Highways and Transportation.
- National Asset Management Steering Group (2000) *International Infrastructure Management Manual*, Wellington: Association of Local Government Engineers New Zealand.
- Phillips, S.J. (1994) 'Development of United Kingdom pavement management system', in TRB, *Third International Conference on Managing Pavements, San Antonio Texas, 22–26 May 1994, Proceedings Volume 1*, Washington, DC: National Academy Press, 227–236.
- Robinson, R. (1998) 'Guidelines for the design and operation of road management systems', *Overseas Road Note 15*, Crowthorne: Transport Research Laboratory for the Department for International Development.
- Robinson, R. (2003) 'Road maintenance management for district engineers', *Overseas Road Note 1*, 3rd edn, Crowthorne: Transport Research Laboratory for the Department for International Development.
- Robinson, R. and May, P.H. (1997) 'Road management systems: guidelines for their specification and selection', *Proceedings of the Institution of Civil Engineers, Transport*, **123** (Feb), 9–16.
- Robinson, R., Danielson, U. and Snaith, M. (1998) *Road Maintenance Management: Concepts and Systems*, Basingstoke: Macmillan.
- Shahin, H.Y. (1994) *Pavement Management for Airports, Roads, and Parking Lots*, New York: Chapman & Hall.
- TRL and DFID (1999) 'A guide to pavement evaluation and maintenance of bitumen-surfaced roads in tropical and sub-tropical countries', *Overseas Road Note 18*, Crowthorne: Transport Research Laboratory.
- TRL and DFID (2003) 'Management of rural road networks', *Overseas Road Note 20*, Crowthorne: Transport Research Laboratory for the Department for International Development.

Maintenance operations

Bent Thagesen

20.1 Introduction

The broad objective of road maintenance is to keep roads as nearly as possible in original condition. However, the resources available for road maintenance are usually limited in industrialized as well as emerging and developing countries. This suggests a more comprehensive statement of the objective of maintenance is needed, such as (Gichaga and Parker 1988):

Objective

To conserve, as nearly as possible, the original designed condition of paved and unpaved roads, and of traffic signs, signals and markings, in a manner most likely to minimize the total cost to society of vehicle operation and accident cost, plus the cost of providing the maintenance itself, under the constraints of severe resource limitations, in respect of skilled manpower, equipment and money, both local and foreign.

In the late 1970s, the foreign aid administrations of France, Germany and the United Kingdom joined forces to produce a handbook with practical guidelines for road maintenance in Africa. Published in 1982, the book soon became widely known and used far beyond Africa. In the early 1990s the handbook was reviewed and made more suitable for a wider audience. This work was undertaken by the Permanent International Association of Road Congresses (PIARC). An English version of the handbook, entitled the *International Road Maintenance Handbook*, was published in 1994 by the Transport Research Laboratory (TRL) in UK in four volumes. The *International Road Maintenance Handbook* forms the main basis for this chapter on maintenance operations. For more detailed information on execution of road maintenance, required personnel, plant, tools, materials and safety equipment, the actual handbook should be consulted.

The
International
Road
Maintenance
Handbook

20.2 Classification of maintenance activities

Maintenance activities may be classified in terms of their operational frequency into

- routine maintenance
- periodic maintenance
- urgent maintenance.

Routine maintenance	<p>Routine maintenance covers activities that must be carried out frequently, that is, once or more per year. They are typically small scale, or simple, and often widely dispersed. There are two types of routine maintenance:</p> <ul style="list-style-type: none"> • <i>Reactive</i> – those whose frequency depends on the volume and intensity of traffic. • <i>Cyclic</i> – those whose need is independent of traffic, and whose frequency relates to rainfall, topography and other local environmental aspects.
Periodic maintenance	<p>Cyclic works can be estimated and planned in advance, for example, vegetation control on shoulders and slopes. Reactive works are more difficult to plan in advance, for example, pot-hole patching. In past documents, the term ‘recurrent maintenance’ was sometimes used – but the term ‘reactive’ is now preferred.</p> <p>Periodic maintenance describes activities that are needed occasionally, that is, after a period of some years. They are usually large scale and require more equipment and skilled labourers than routine maintenance activities.</p>
Urgent maintenance	<p>Urgent maintenance comprises emergency repair required by flood damage, earth slips, overturned trees, etc. In some countries, winter maintenance would be included under this heading. However, in those countries where severe winters are the norm, winter maintenance would be considered a routine (reactive) activity.</p>
Contents	<p>Maintenance activities and the defects that they treat are discussed under the following headings:</p> <ul style="list-style-type: none"> • Asphalt pavements • Unpaved roads • Roadside areas • Drainage systems • Traffic control devices.

The book does not cover winter maintenance, cement concrete pavements, nor the maintenance of bridges.

20.3 Safety

Precautions	<p>Before starting any maintenance work, precautions should be taken to ensure the safety of both road workers and road users.</p>
Safety vests	<p>Yellow or orange coloured safety vests should be worn by all workers. Vehicles and equipment should also be painted yellow or orange and carry high-visibility markers front and rear. The vehicles should work with headlights switched on and, where possible, carry yellow flashing lights or flags.</p>
Advanced warning	<p>Traffic signs should be placed ahead of the worksite in both traffic directions, to give advanced warning of danger. Traffic cones should be placed along the length of the roadworks to protect the site from traffic. A sign should be placed at the end of the roadworks, in both directions, indicating the end of restrictions.</p>
Traffic control	<p>If the traffic can only pass the worksite in one lane, a barrier should be placed at each end, and traffic controllers should stand next to the barriers and operate a reversible ‘stop/go’ sign. On long work sections and sections with restricted visibility, hand-held radio sets, or intermediate traffic controllers, are needed to transfer</p>

the ‘stop/go’ instructions to the traffic controllers. On large worksites, portable traffic lights may be more expedient. In some cases, it may be necessary to close the road temporary and divert the traffic. Simpler systems of traffic control can be used on roads with low traffic volumes. The type, number and exact location of the different traffic signs and traffic barriers depend on the road class and the type, extent and duration of the roadworks. It is important that all temporary warning signs are removed immediately after the roadworks have been completed.

20.4 Asphalt pavements

20.4.1 Activities

Maintenance of asphalt pavements consists of

- Routine activities
 - sanding;
 - local sealing;
 - crack sealing;
 - filling depressions;
 - surface patching;
 - base patching.
- Periodic activities
 - surface dressing;
 - fog spray and slurry seal;
 - asphalt overlays;
 - pavement reconstruction.

20.4.2 Sanding

Bleeding is a migration of bitumen to the surface of the pavement. It is usually caused by too much binder in the surfacing or by an unsuitable binder. Bleeding reduces surface friction and causes the road surface to become slippery.

Bleeding

Sanding can be used to treat bleeding. A thin layer of sand is spread over the surface. The sand is scattered with a shovel and spread out with a broom. Whenever possible, the sand should be coarse grained with a particle size 0–5 mm.

Treatment

20.4.3 Local sealing

Local sealing is used to repair more serious bleeding, for sealing local cracks and as a final treatment when carrying out base patching.

Purpose

Cracking may occur in the pavement structure, or in the surfacing only. Longitudinal cracking is often localized along the wheel tracks or along the edges of the pavement. Transverse cracking appears across the whole, or part of the cross section. Serious cracking will develop into ‘mesh’ cracking, dividing the pavement

Cracking

surface into isolated polygons of different sizes. Mesh cracking is also called ‘block’, ‘crocodile’ or ‘alligator’ cracking. Fine longitudinal cracks in the surfacing may be caused by ageing of the pavement. Deep cracks are usually due to insufficient pavement thickness for the traffic being carried, or poor drainage that allows water to penetrate and weaken the pavement and the subsoil. Isolated transverse cracks are typically caused by shrinkage of a cement-treated base. Cracks in asphalt pavements may also be due to poor quality materials and poor workmanship. In climates with wide temperature variations, cracks may develop if the bitumen in the asphalt is of a type that becomes brittle at low temperatures. Cracking allows rainwater to seep into the road, and can be the beginning of a local or more general destruction of the pavement. Only surface cracking and slight cracking in the pavement structure should be repaired with local sealing. Severe cracking in the pavement structure requires base patching.

Treatment

Local sealing is the application of a surface dressing over a local area. First, the area is swept clean; then about 1.5 kg of emulsified bitumen, or 1 kg cutback bitumen is spread per square metre of surface using a pressure distributor with spray lance, or a watering can. After the binder has been applied, aggregate is distributed with a shovel. The aggregate should normally be coarse sand. However, when finishing base patching, the aggregate should normally be 6–10 mm chippings. Chippings should be rolled into the bitumen using a small roller or truck tyres.

20.4.4 Crack sealing

Closely spaced cracks may be filled with an asphalt slurry rather than by local sealing. A slurry is produced by mixing 20 litres of coarse sand with 6 litres of emulsified bitumen. Again, the area should be swept clean. The slurry is spread out in a thin layer with a wooden board fitted with a handle. An alternative method of sealing isolated cracks is to fill the cracks with hot cutback bitumen.

20.4.5 Filling depressions

Purpose

This treatment is applied to deal with slight depressions and rutting, slight edge subsidence and small surface irregularities due to shoving. Deep depressions and rutting, deep edge subsidence and large surface irregularities usually require base patching.

Depressions

Depressions are small hollows of limited size in the pavement. Depressions usually develop from random local defects, such as low compaction or insufficient pavement strength. Depressions increase the roughness of the pavement and deep depressions may be a safety hazard to the traffic.

Rutting

Rutting, illustrated in Figure 20.1, is longitudinal subsidence localized in the wheeltracks caused by vehicles. The main causes of rutting are inadequate stability of the asphalt material in the surfacing, inadequate compaction of the pavement and insufficient pavement strength. Water may accumulate in ruts during rain and expose the road users to aquaplaning. In cold climates, freezing water in ruts can be very dangerous. If water is able to penetrate the surfacing in the ruts, it may lead to cracking and breaking up of the pavement.

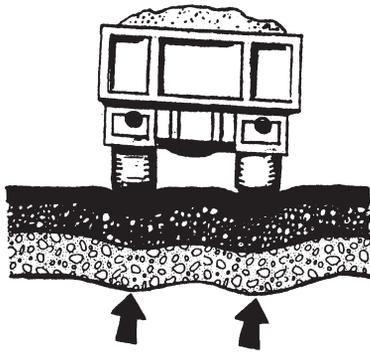


Figure 20.1 Rutting.

Source: *International Road Maintenance Handbook* (PIARC 1994).

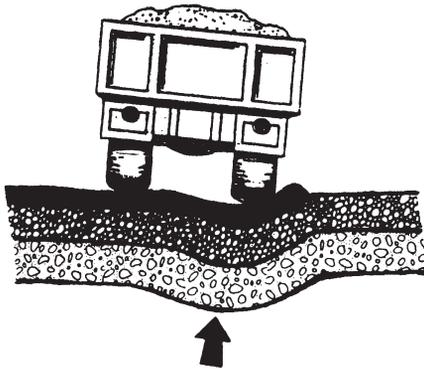


Figure 20.2 Edge subsidence.

Source: *International Road Maintenance Handbook* (PIARC 1994).

Edge subsidence, shown in Figure 20.2, occurs where the pavement borders unsurfaced shoulders. The main causes are inadequate, or badly maintained shoulders and poor drainage that allows water to penetrate into the pavement structure below the edges. Another cause may be a narrow roadway that forces the traffic to drive near the edge. Edge subsidence can encroach to affect significant areas of the pavement.

Edge
subsidence

Shoving, shown in Figure 20.3, is a horizontal displacement of the surfacing. It often occurs on either side of the wheeltracks and also near crossroads where vehicles brake and accelerate. The main causes of shoving are lack of stability and insufficient compaction of the surfacing and/or base. Shoving increases the roughness of the road and may lead to disintegration of the pavement.

Shoving

The depressions must be swept clean and dry. A tack coat of hot cutback bitumen is spread at a rate of about 0.5 kg/m^2 using a pressure distributor with a spray lance or a watering can. Cold mix asphalt is placed in the depression and compacted with a hand rammer or a roller. To prevent penetration of water, local sealing of the repair is recommended.

Filling
depressions

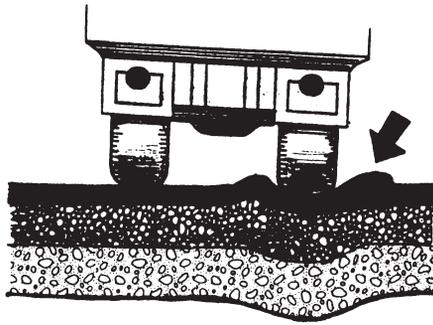


Figure 20.3 Shoving.

Source: *International Road Maintenance Handbook* (PIARC 1994).

20.4.6 Surface patching

Local aggregate loss

Surface patching is used to repair local aggregate loss. Loss of aggregate from the surface of a premixed asphalt pavement (fretting) is usually due to poor premix quality or poor workmanship. Loss of chippings from a surface dressing (stripping) may be caused by insufficient binder, use of dirty chippings, or insufficient penetration of the chippings into the binder. Loss of aggregate from a surfacing will cause slow disintegration of the layer.

Surface patching

Surface patching may be carried out using the same treatment as for local sealing (surface dressing). An alternative method is to use asphalt premix. Cold premix is easiest to handle. After spraying cutback bitumen on the surface at a rate of 0.5 kg/m^2 , the cold mix is distributed evenly over the area and compacted with a small roller or a hand rammer.

20.4.7 Base patching

Purpose

Base patching is used for local restoration of the pavement structure, including repair of severe mesh cracking, deep rutting and depressions, deep edge subsidence and rutting, broken edges, pot-holes and severe shoving.

Broken edges

Broken edges (see Figure 20.4) may result from low shoulders, penetration of water along the edges, insufficient compaction of the pavement edges and narrow pavements. Broken edges are self-perpetuating defects. They may be a safety hazard to traffic.

Pot-holes

Pot-holes are local holes in the pavement where materials have been removed by the action of traffic and water. Pot-holes usually develop in areas showing cracks, deformations or aggregate loss. In new surface dressings, pot-holes can develop from random local defects in the surface or base. Pot-holes drastically increase the roughness of the road and deep pot-holes are a safety hazard to road users.

Base patching

First, all loose material is removed from the damaged area; the depth of the hole is then increased until firm material is found. The sides of the hole are cut back to vertical and the bottom of the hole is trimmed and compacted flat and parallel with the road surface (see Figure 20.5). Any water in the hole must be removed. The

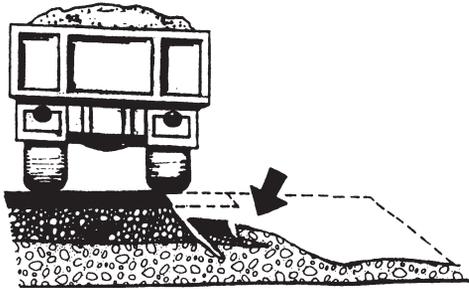


Figure 20.4 Broken edges.

Source: *International Road Maintenance Handbook* (PIARC 1994).

prepared hole is filled with material of the same quality as that of the pavement to be repaired, or with an asphalt mix. The material should be placed in the hole in one or more layers and compacted with a hand rammer or a roller, depending on the size of the patch.

20.4.8 Surface dressing

Surface dressing of the complete width of the pavement is appropriate if large areas are damaged by bleeding, cracks, slight depressions or aggregate loss. It is also the remedy for streaking and glazing.

Streaking is loss of aggregate from a surface dressing in streaks running parallel to the pavement centre line. Streaking is normally caused by faulty operation of the spraying equipment.

Glazing is due to embedment of chippings in the surface giving a smooth, shiny appearance. The main causes are wear (not removal of the surface chippings) and embedment of the chippings into the base.

Surface dressing may be carried out using equipment-based methods as well as by labour-based methods. It is possible to make the quality of labour-based work as good as that using mechanized methods if the work is properly managed. Information about design and execution of surface dressing is given in Chapter 14. Before applying a surface dressing over large areas, it will normally be necessary to carry out some patching work, and repairs to shoulders and the drainage system. Surface dressing should not be used to repair severe cracking that is due to insufficient pavement strength or ageing of the base. Instead, the pavement needs strengthening with an asphalt overlay or complete reconstruction.

20.4.9 Fog spray and slurry seal

Fog spray and slurry seal are options that, in certain circumstances, may be more appropriate than a surface dressing. A fog spray is a very light film of emulsion, which is sprayed onto the surface of an old, lean asphalt surface in order to hold stone particles in place that otherwise would be picked off by the traffic.

Purpose

Streaking

Glazing

Surface dressing

Fog spray

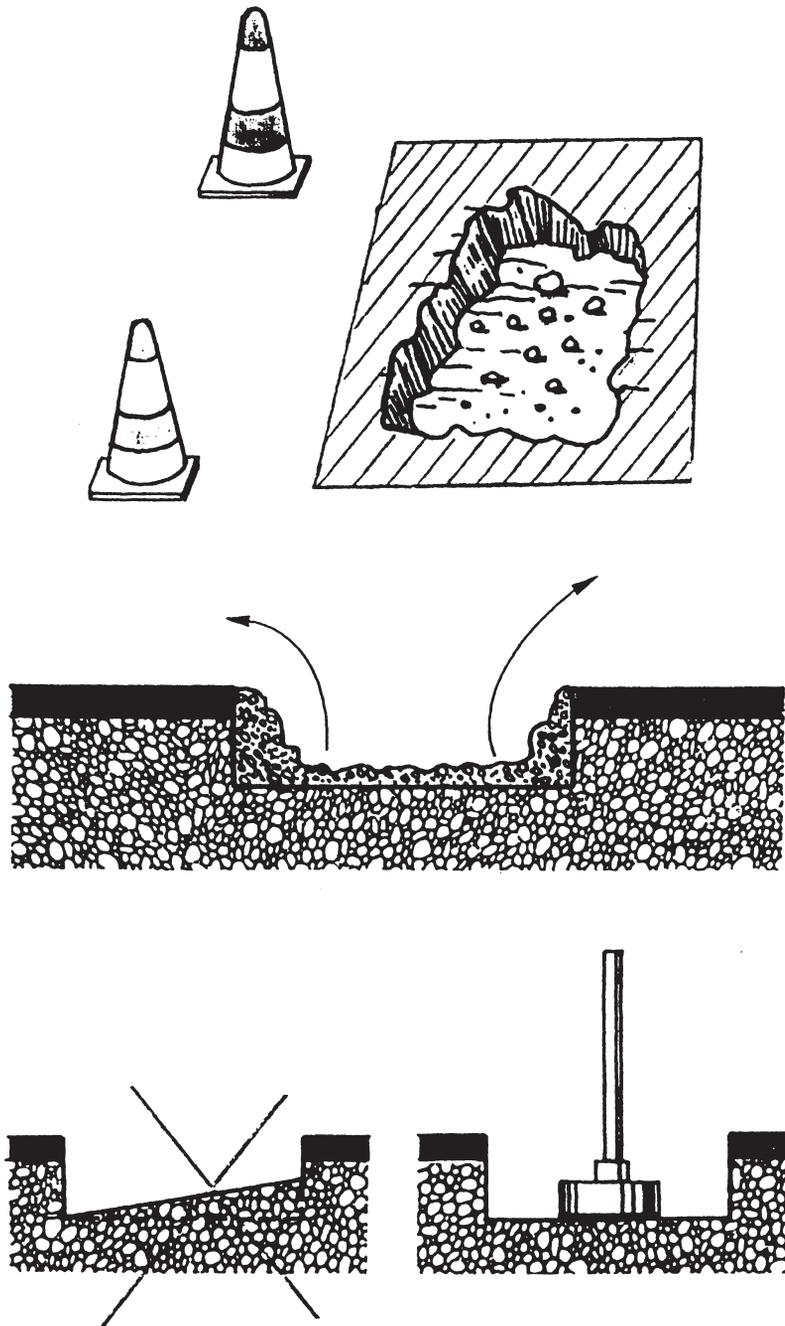


Figure 20.5 Patching of pot-holes.

Source: *International Road Maintenance Handbook* (PIARC 1994).

Slurry seal

A slurry seal is the application of a mixture of fine aggregate and emulsified bitumen over the full width of the pavement, similar to the crack sealing described earlier. The asphalt slurry penetrates and seals surface voids and cracks very effectively. Slurry seals can be prepared in a concrete mixer using a slow-breaking

emulsion and then spread on the road by hand. However, the normal technique is to use a mechanized mixer and spreader unit, which enable faster breaking emulsions to be used.

20.4.10 Asphalt overlay

An asphalt overlay is an application of hot, premixed asphalt. Overlays are used to strengthen old pavements or pavements with insufficient thickness. Serious bleeding and severe shoving may also warrant an asphalt overlay. Repairs to pot-holes and edge damage, and restoration of shoulders and the drainage system need to be undertaken before strengthening. It may also be necessary to repair rutting and to fill depressions. Strengthening should always be designed properly on the basis of a thorough examination of the existing pavement, as described in Chapter 15. Information about mix design and construction of hot mixed asphalt pavements is given in Chapter 14.

20.4.11 Pavement reconstruction

Reconstruction involves strengthening by multiple overlays or removing some of the existing layers prior to providing new pavement materials. Recycling of the base and surfacing is being used increasingly for cost and environmental reasons, but is only appropriate where the materials being recycled are reasonably homogeneous. Reconstruction is used mainly when complete failure of the construction has occurred, but is also appropriate when this solution is cheaper than an overlay made entirely of asphalt mixture. Most road administrations classify reconstruction as a construction activity rather than a periodic maintenance operation.

20.4.12 Identifying the causes of deterioration

Asphalt roads will generally deteriorate either by cracking or rutting, but other types of surface defect also occur, such as bleeding, aggregate loss and streaking. It is important that the initial form of deterioration and its cause is identified, because this determines the type of maintenance that is most appropriate. Treating the symptoms of pavement defects rather than the causes will prove unsatisfactory. To help identify the causes of deterioration, *Overseas Road Note 18* (TRL and DFID 1999) subdivides cracking and rutting into six categories:

- rutting without shoving;
- rutting with shoving;
- wheelpath cracking in premixed asphalt surfacing;
- wheelpath cracking in surface dressing;
- non-wheelpath cracking in premixed asphalt surfacing;
- non-wheelpath cracking in surface dressing.

The probable cause or causes of the deterioration are established using a flow chart, such as that in Figure 20.6.

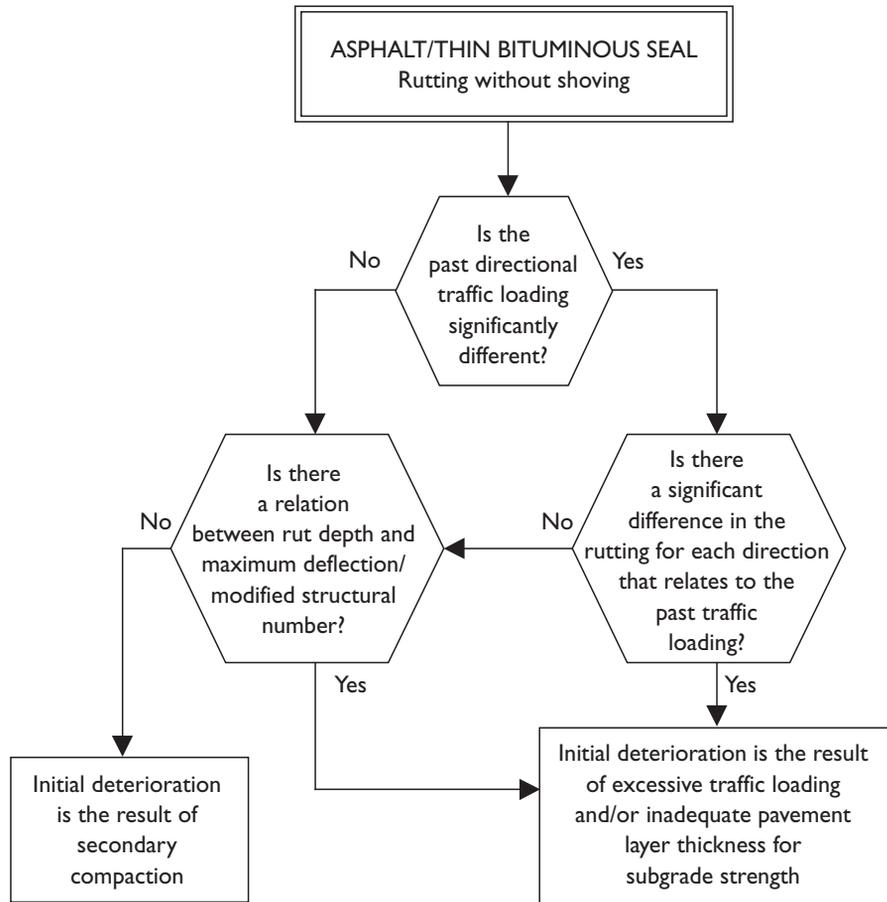


Figure 20.6 Initial deterioration – rutting without shoving.

Adapted from: *Overseas Road Note 18* (TRL and DFID 1999).

20.5 Unpaved roads

20.5.1 Road types and activities

Unpaved roads are either earth roads constructed from the natural soil found on the route, or gravel roads surfaced with a layer of gravel that is stronger than the natural soil. Maintenance of unpaved roads includes

- Routine activities
 - grading;
 - labour-based minor reshaping;
 - dragging;
 - patching.
- Periodic activities
 - regravelling.

20.5.2 Grading

Grading is used to remove poor shape, ruts, pot-holes, corrugations and erosion gullies in the road.

Poor shape refers to a condition where depressions and lack of crossfall prevent rainwater from draining easily from the surface of the road. Ruts and pot-holes are types of defects similar to those in asphalt pavements. All these defects may be caused by inadequate compaction or instability of the materials and, for gravel roads, by insufficient thickness of the gravel layer. The defects are self-perpetuating as they will intercept rainwater and allow the ingress of water and weaken the road.

Corrugations are transverse waves on the road surface, caused by the bouncing of the wheels of vehicles (Heath and Robinson 1980). Corrugations are generally quite hard and about 25–40 mm in amplitude. The wavelength is usually between 300 and 900 mm and fairly uniform at any particular site. Corrugations only occur on a range of more granular soils. Driving on corrugated roads can be hazardous, is extremely unpleasant and increases the wear on vehicles.

Erosion gullies are channels in the road surface caused by rain, and occurring mainly on steep grades. Once erosion gullies have formed, they may quickly increase in size and destroy the road section concerned.

Grading is carried out by self-propelled or towed graders. A grader restores the crossfall and removes surface irregularities by returning materials from the sides and shoulders towards the centre of the road. In advance of grading, large pot-holes and depressions should be patched, areas of standing water should be drained and the side ditches must be cleaned and reshaped which, depending on their cross-section, can also be undertaken with a grader. Patching of unpaved roads and cleaning and reshaping of ditches are described later in this chapter.

The grader should work on one side of the road at a time in passes of about 200 m long. Where the defects are limited, two or four passes of the grader may be enough to reshape the road. If the defects are more severe, additional passes are needed. In cases where the road is badly damaged, it may be necessary to scarify the existing surface to the bottom of any defects and loosen the material before reshaping. Work should be completed on one side of the road at a time. Shoulders are treated as part of the running surface. Whenever possible, rollers should be used to compact the road after grading. Work is best scheduled to follow a period of rain, as the moisture in the material will help compaction whether rollers are used or the compaction is left to traffic.

20.5.3 Labour-based minor reshaping

Labour-based minor reshaping provides an alternative to grading when maintaining roads with limited defects and very low traffic volumes. The labourer trims the surfacing material with a pickaxe, hoe or mattock, and rakes it to form the specified crossfall. On gravel roads, any local depressions and pot-holes are filled with fresh gravel from stockpiles placed along the road. Whenever possible, the loose material should be compacted with a roller or a hand rammer.

20.5.4 Dragging

Dragging is a simple method used to smooth out minor defects in the road surface. It is particularly useful as a frequent operation on roads with a sandy or loose

Purpose

Poor shape,
ruts and
pot-holes

Corrugations

Erosion gullies

Grading

Grader
operation

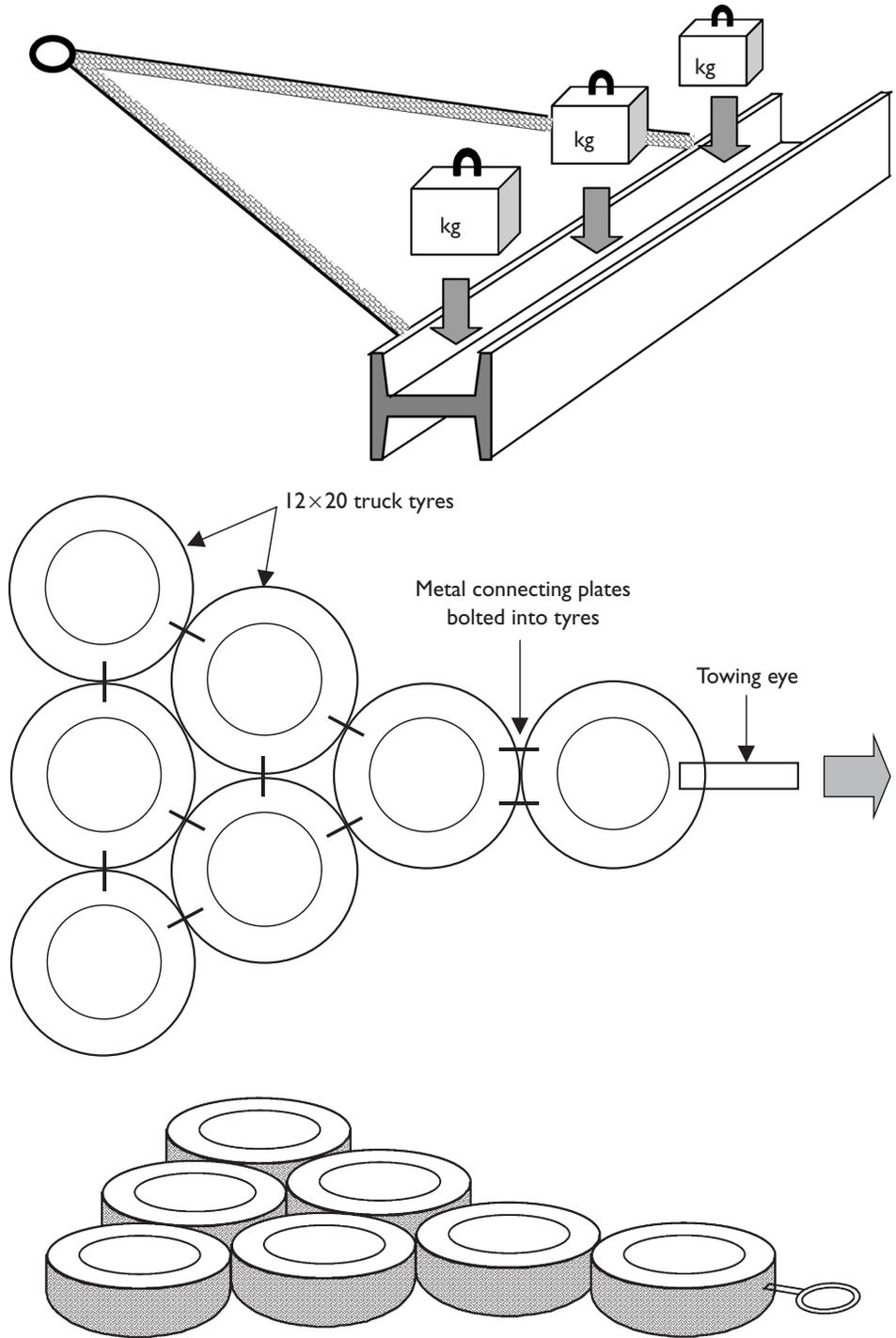


Figure 20.7 Beam drag and tyre drag.
Adapted from: TRRL Overseas Unit (1985).

surface. The operation is usually carried out by towing a specially made drag behind an agricultural tractor. Figure 20.7 shows two different types of drag, one made from old truck or tractor tyres cut in half around their circumference, and the other made from a steel rail. On low volume roads, frequent dragging may be used to reduce the need to grade the road. However, dragging will not remove corrugations once they have formed, nor will it restore crossfall and lost material. After a number of dragging operations it is usually necessary to grade the road in order to return materials from the sides and shoulders towards the centre of the road.

20.5.5 Patching

Patching of unpaved roads consists of adding new surfacing material over a relatively small area. Patching may be used to repair deep surface irregularities and to make good local areas with inadequate crossfall. Large-scale patching of gravel roads is called spot regravelling and considered to be a periodic maintenance activity. Patching is not a suitable method for repairing corrugations.

The quality of the material used for patching should be at least as good as the quality of the material already surfacing the road. Topsoil containing vegetable matter should never be used for patching.

Large pot-holes are trimmed and their sides cut back to vertical. The hole is then filled with new material in layers of about 100mm thickness. Each layer is compacted with a hand rammer or a roller, depending on the area of the hole. Whenever possible, dry material should be sprinkled with water to assist the compaction. When labour-based maintenance is used, gravel for patching should be stockpiled every 100–200m along the road without blocking the road or the side drains. This is most conveniently done when the periodic maintenance is carried out. A useful place for stockpiling is downhill of turnout drains.

20.5.6 Regravelling

Regravelling is used on gravel roads to replace lost surfacing materials. Regravelling is also used to correct severe surface irregularities and road sections lacking adequate crossfall.

Material loss is caused by erosive wear due to traffic and water. Dry, loose material wears off quicker than cohesive materials and materials containing a little moisture. For gravel roads, the material loss inevitably results in a continuously decreasing thickness of the gravel layer. Regravelling is needed before the thickness of the existing gravel layer is reduced and provides insufficient coverage over the subgrade. If a road is not regravelled in time, ruts and pot-holes will quickly appear, and the road may soon become impassable.

Before regravelling is carried out, it is important to make any necessary repair or improvements to the crossfall of the road and the drainage system. Gravel of a quality that meets existing specifications should be stockpiled at the quarry or borrow pit. The regravelling operation may be executed by equipment-based or labour-based methods.

The road must first be graded and compacted to provide a firm, regular surface on which to work. New gravel is transported by trucks from the quarry or borrow pit to the worksite. Dumping of the gravel should start from the far end of the worksite,

Purpose

Materials

Treatment

Purpose

Material loss

Treatment

Equipment-based approach

and the material should be placed on one side of the road only and with the spacing necessary to give the intended thickness of gravel over the complete width of the road. The material is now spread across the road using the grader. Water is added from a water tanker. When a correct camber has been achieved, the gravel is compacted with a roller.

Labour-based approach

When adopting a labour-based approach, the existing surface is first reshaped to correct crossfall without using silty soil from the ditch or topsoil containing vegetable matter. The material is compacted with hand rammers or a roller. New gravel is usually transported on tractor-drawn trailers. In the borrow pit, the gravel should be excavated and stockpiled in a way that allows the trailers to be easily loaded by hand, as in Figure 20.8. Each tractor should work with at least two trailers to maximize use of the tractors. On the worksite, the trailer is unloaded within a rectangle marked with pegs set at the finished road level. The gravel is spread to correct level and crossfall using shovels, hoes and rakes. Whenever possible, the gravel should be compacted with a roller and dry material should be watered prior to compaction.

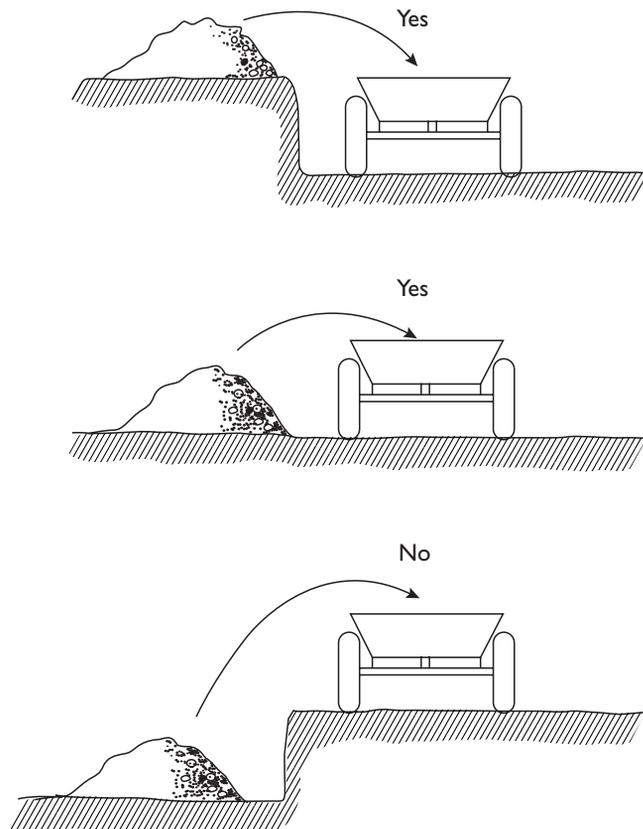


Figure 20.8 Hand-loading of trailers.

Source: *International Road Maintenance Handbook* (PIARC 1994).

20.6 Roadside areas

20.6.1 Features

Roadside areas consist of shoulders, slopes and other surface areas within the road margin. Paved shoulders and lay-bys are treated as pavements. Most maintenance activities for roadside areas can be carried out either by labour or machinery.

20.6.2 Shoulders

20.6.2.1 Activities

The following maintenance activities are needed on shoulders:

- Routine activities
 - removing obstructions;
 - reshaping shoulders;
 - vegetation control.
- Periodic activities
 - adding shoulder materials.

Note that the shoulders of unpaved roads become indistinguishable from the main carriageway after construction, and will be maintained along with the carriageway. The maintenance of paved shoulders is very similar to that for the paved carriageway.

20.6.2.2 Removing obstructions

Shoulders may be obstructed by fallen rock, trees or branches, soil heaps, windblown material, abandoned vehicles and debris. The obstructions are a hazard to road users and may prevent flow of water from the road to the ditches. They should be removed and disposed of to a safe location.

20.6.2.3 Reshaping shoulders

The surface of shoulders may become higher than the carriageway and the surface mis-shapen because of traffic on the shoulders, accumulation of loose material from the carriageway, accumulation of soil slipping down from cutting slopes, or foreign material trapped in the vegetation on the shoulders. High or mis-shapen shoulders are a hazard to the traffic and may cause water to accumulate along the edges of the road and that will weaken both pavement and shoulders.

High/
mis-shapen
shoulders

High or mis-shapen shoulders are repaired by grading, or by labour-based reshaping, to correct the crossfall. When grading shoulders on paved roads, care should be taken not to damage the pavement edges with the grader blade. Excess material from the shoulders should be removed and not deposited on the road, or in the ditch.

Reshaping

20.6.2.4 *Vegetation control*

High
vegetation

Grass, weeds and bushes, allowed to grow unchecked on the shoulder, may cause water and silt to accumulate at the edge of the road and may, at worst, impair sight distance for road users. In the dry season, unchecked vegetation may constitute a fire hazard.

Vegetation
control

Vegetation should be cut at least once a year. Most of the vegetation may be trimmed by an agricultural tractor equipped with a mower, or using a hand-guided power mower. As an alternative to mowing, and in areas inaccessible to power equipment, the vegetation should be trimmed by hand using sickles, scythes, slashers, bush knives and similar hand tools.

Herbicides
and burning

It is not recommended that chemical methods be used to control roadside vegetation. Herbicides can be dangerous to humans and fauna, and can pollute crops and water courses. Similarly, roadside vegetation should not be controlled by burning, as the flame and smoke blowing across the road are dangerous for the traffic. Fire can spread and destroy valuable vegetation and crops, and can be dangerous for animal life.

20.6.2.5 *Shoulder regravelling*

Low/
mis-shapen
shoulders

Additional shoulder material is required when the shoulder becomes too low, or is badly mis-shapen with ruts and depressions. This is an activity similar to regravelling unpaved roads. This degree of shoulder degradation may result from over-running by traffic, erosion by water, or settlement. A low shoulder gives inadequate support for the pavement. Water can collect in ruts and depressions and soften the shoulder and pavement edge. The pavement edge can also break away when vehicles run over it. Furthermore, the danger of accidents is increased.

Treatment

When using equipment, the existing surface of the shoulder is scarified with the tines of a grader. New material is offloaded onto the shoulder and shaped to correct crossfall and to a level slightly above the final level. On paved roads, care must be taken not to damage the edge of the pavement with the grader blade. The shoulder is then compacted with a roller. When using labour, the existing surface of the shoulder should be loosened with a pickaxe or mattock. New material is added and shaped using shovels, hoes and rakes. The shoulder is compacted using hand rammers or a manually propelled roller. In both cases, dry material should be watered to assist compaction whenever possible.

20.6.3 *Slopes*

20.6.3.1 *Activities*

Slope maintenance includes

- routine activity – vegetation control;
- periodic activity – erosion control;
- periodic or urgent activity – slip repair.

20.6.3.2 *Vegetation control*

Overgrown vegetation, due to insufficient grass cutting, bush clearing and tree trimming, may reduce visibility for the road users. Tilting trees can fall and block the

ditch and road. Overgrown vegetation is a fire hazard in the dry season. The activities necessary to control vegetation on slopes are the same as for shoulders, but the required frequency is usually less.

20.6.3.3 Erosion control

Erosion of slopes is usually caused by rainwater concentration on the slopes, or lack of vegetation cover. If neglected, the erosion may develop into deep ravines and cause the slope material to slide downhill. Eroded material may block the roadside ditch and shoulder.

Erosion

Eroded areas may be repaired with grass turfing or seeding. These methods will only be successful if the climate and soil conditions are favourable. Small, important areas at bridges and culverts may be protected with riprap (stone pitching) provided that the slope is not steeper than 1 vertical to 1.5 horizontal. Any rough stone can be used for riprap. The size should be as uniform as possible and the individual stones should weigh 10–20 kg. Heavier stones are preferred if the riprap is not to be grouted.

Erosion repair

To prevent future erosion of a cutting slope, a soil berm or a cut-off drain can be built along the top of the slope to prevent surface water from flowing down and eroding the face of the cutting. Berms and cut-off drains should be constructed to lead all water to a safe discharge location. For embankments, a kerb or channel drain may be built at the back of the shoulder. The kerb may be of dressed stone or concrete. If the gradient falls throughout the embankment, the kerb or channel drain may discharge at the transition between fill and cut. If there is a low point on the embankment, a chute or a cascade will be required to discharge the water safely down the slope. A chute is an inclined pipe-drain constructed in or on a slope. A cascade is a channel with a series of steps, to take water down a steep slope, and dissipate some of the energy of flow.

Slope protection

20.6.3.4 Slip repair

Slips or slope instability are usually caused by adverse ground conditions, or water, or both. The slope may be too steep for its height and soil type. Water may be penetrating the slope from above, or there may be ground water pressure.

Slips

All slipped soil should be excavated from the road, shoulder and drain, and removed to a suitable dump site. However, slip repair is a dangerous activity and it is often advisable to reduce the slope angle before clearing the material that has slipped. Where a cutting face has slipped, an option is to reduce the angle of at least the upper part of the slope. A gabion retaining wall may be used to stabilize the base of the slope. A gabion is a stone-filled steel mesh cage. A gabion retaining wall is free draining and sufficiently flexible to allow for further small slip movements. Where the foundation is stable, an option is to construct a masonry or concrete retaining wall.

Slip repair

20.7 Drainage systems

20.7.1 Activities

Maintenance of the drainage system includes the activities listed in Table 20.1. Most drainage maintenance can be achieved by labour-based methods.

Table 20.1 Activities included in the maintenance of the drainage system

<i>Elements</i>	<i>Routine activities</i>	<i>Periodic activities</i>
Drains	<ul style="list-style-type: none"> • Clearing and cleaning • Reshaping and deepening • Minor erosion control 	<ul style="list-style-type: none"> • Providing new turnouts • Major erosion control
Culverts	<ul style="list-style-type: none"> • Clearing and cleaning • Erosion repair • Cracking repair • Walls and apron repair 	<ul style="list-style-type: none"> • Repair of invert • Construction of outfall basin • Reconstruction
Fords, drifts and causeways	<ul style="list-style-type: none"> • Clearing • Repair • Replacing of guide posts 	
Drainage pipes and manholes	<ul style="list-style-type: none"> • Clearing of pipe and manhole • Replacing of manhole and grating • Clearing of manhole area • Cleaning of catchpit sump 	<ul style="list-style-type: none"> • Re-lay drainage pipe

20.7.2 Drains

20.7.2.1 Clearing and cleaning

The object of clearing and cleaning drains is to remove all obstructions that could possibly interfere with water flow. Obstructions can include loose silt, sand, gravel, boulders, weeds, bushes and fallen trees. Obstructions should be removed and disposed of well away from the roadside, so they will not fall or wash back into the drain. Where a grassed ditch has the correct depth and profile, with no erosion, the grass should just be cut short. The grass roots will help to bind the surface together.

20.7.2.2 Reshaping and deepening

Destroyed cross-section, silting and ponding
Reshaping and deepening

The cross-section may be destroyed due to vehicular or animal traffic, or due to collapse of the ditch sides. Silting and ponding (standing water) occurs where the invert slope of the drain is so flat that the water cannot flow at sufficient speed. Ponding may also be caused by a drain cross-section that is too small.

A grader can be used to reshape and deepen V-shaped ditches. Trapezoidal side drains can be maintained with a tractor-mounted back-hoe, or can be maintained by labour using the hoe, shovel and ditch template. In flat areas, the gradient of the drain should be checked carefully. Excavated material must be removed and spread well clear of the drain so that it cannot later fall or wash back into the drain. Material from the ditches should not normally be spread on the running surface of the road as its high content of fine-grained material makes it unsuitable as surfacing material.

20.7.2.3 Minor erosion control

Erosion

Erosion of drains occurs where velocity of flow of the water is too great. Erosion is often seen in drains laid at a steep gradient, on sharp bends without erosion protection, and at drain outfalls.

Unlined drains that have been damaged by scour of the invert and sides may be repaired by filling with soil and turfing, provided that the climatic conditions are favourable. To avoid future erosion, simple scour checks may be constructed of wood pegs or stones. All scour checks should have an apron downstream built of stones or grass turfs pinned to the ditch invert with wooden pegs. Lined drains, where the linings have been damaged, should be repaired as soon as possible.

Erosion repair

To avoid future erosion, drain sections with a steep gradient may be realigned to follow contour lines more closely. Sections with sharp bends may be re-laid to smooth curves, or special precast curved sections may be installed. Drain outfalls may be extended to reduce the speed of the water when leaving the ditch. The gradient should ideally be between 2 and 5 per cent. Areas downstream from outfalls may be protected by turfing or stone pitching.

Preventive measures

20.7.2.4 Addition of turnouts

Extra turnouts may be required where water ponds in the side drains, where drains are hydraulically overloaded, or where erosion occurs. Turnouts are desirable at spacings as close as 20 m on some gradients. If water cannot be discharged from a drain over a distance of 200 m, it may be necessary to line the drain or construct an extra culvert.

20.7.2.5 Major erosion control

Major erosion control includes lining of ditches, where they are frequently damaged, and construction of chutes or cascades in places where large volumes of water have to be taken down slopes.

Treatments

Ditch linings may be constructed with masonry stone, brick, precast concrete tiles or precast drain units. It is important that the earth profile is compacted carefully before the stone or concrete panels are bedded and jointed with cement mortar, otherwise the masonry may lift off during heavy rain. Construction of chutes is discussed in Chapter 10.

Lining of ditches

A cascade may be constructed of stone masonry, brick or concrete. Instead of solving erosion problems, cascades often create problems themselves because of inadequate design and construction. A common defect is that the steps are built with a vertical face greater than the horizontal face.

Cascades

20.7.3 Culverts

20.7.3.1 Cleaning and clearing

Culverts may be blocked by sand or floating vegetation and debris. This happens particularly when the invert slope of the culvert is too flat and when the invert level is lower than the bottom of the downstream ditch. These are often construction or design faults. The result of a blockage is that the intended waterway opening is reduced to the extent that floodwater cannot flow. The floodwater will back-up or pond on the upstream side of the culvert and may eventually overflow the road embankment. The road is then in danger of being severely eroded and washed away.

Blockage

Clearing and cleaning Material and debris blocking culverts must be removed and dumped where they cannot impede water flow. In addition, the upstream approaches and the downstream area must be freed of any obstructions. If floating debris is a recurrent problem, the provision of a debris rack should be considered. Cleaning of culverts with openings smaller than 1 m is a particular problem. Sometimes small culverts can be cleaned by pulling a cable or rope through, to which is attached a bucket.

20.7.3.2 *Erosion repair*

Erosion A commonly occurring defect is erosion of the stream bed at the culvert outlet. This happens when the culvert invert has been constructed with excessive gradient so that the velocity of water flow is too great, or when the culvert invert has been constructed with too flat a slope and an excessive drop at the outfall. These defects are normally design or construction faults and, if neglected, the stream bed is washed away. This may cause collapse of the downstream headwalls and wingwalls, and even sections of the culvert and road embankment.

Repair Where only light erosion has taken place, the eroded area should be filled with stone blocks of about 300 mm size to produce a rough energy dissipator. The block pitching should preferably extend beyond the eroded area. Where stone is not available, logs or jute sacks filled, with soil and cement, can be used.

20.7.3.3 *Cracking repair*

Cracking in headwalls, wingwalls and the main structure is usually due to settlement of the foundation soil below the culvert. If the settlement is small and limited in extent, only minor cracking will result and this will have little effect on the functioning of the culvert. However, cracks should be repaired as soon as possible to reduce water penetration of the foundation and further settlement. To repair cracks, first they are cleaned carefully with a brush and water, and all loose material is removed. The cracks are then wetted and filled with cement mortar.

20.7.3.4 *Walls and apron repair*

Erosion or settlement damage to headwalls, wingwalls or aprons of masonry or brick headwalls should be repaired as quickly and effectively as possible. The settled or damaged section should be removed and the underlying soil compacted. Then the walls and/or apron should then be rebuilt using materials similar to the original. When the repaired walls have recovered sufficient strength, the backfill should be restored.

20.7.3.5 *Repair of invert*

Corrosion The inverts of steel culverts are often rust corroded because the protective galvanizing has worn away. Old steel culverts are particularly prone to this. Repair should be carried out as soon as the surface starts rusting. If neglected, the damage may accelerate and eventually cause collapse of the structure.

Repair Inverts are best repaired in the dry season. Steel culverts with superficial rust damage are cleaned with a steel brush. After cleaning, a thick coat of bitumen is

applied over the lower half of the culvert. If the invert is seriously damaged, a plain cement concrete slab should be poured over the complete length of the invert. The inlet and outlet aprons should also be concreted and adjusted to the new levels of the culvert invert.

20.7.3.6 Construction of outfall basin

If the outfall of a culvert suffers from continual erosion, an outfall basin (catchpit) should be constructed, as in Figure 20.9. The outfall basin will reduce the energy of the water and decrease the risk of downstream erosion. An outfall basin may be constructed of stone masonry or concrete.

20.7.3.7 Reconstruction

Culverts that have collapsed or are beyond repair should be reconstructed. Where erosion problems exist because of excessive water flow, additional culverts may be needed to discharge excess flow to a less overloaded receptor. Consideration should also be given to replacement of culverts with diameter of less than about 600 mm with pipes of a larger size. Reconstruction may be required if the culvert and outfall repeatedly silt-up because the culvert has been constructed too low. The road level may then need to be raised for a distance to accommodate raising of the culvert.

20.7.4 Fords, drifts and causeways

Water will deposit soil and debris on fords, drifts (paved fords) and causeways (vented drifts) from time to time. This creates a hazard to the traffic and exposes the crossing to erosion. Material and debris covering fords, drifts and causeways should be removed and disposed of downstream, well clear of the crossing.

Drifts and causeways are often damaged by cracks, settlements and erosion due to water turbulence. These defects should be repaired as quickly as possible. If not, the defects may spread and cause undermining and disintegration of the drift slab or the

Clearing

Repair

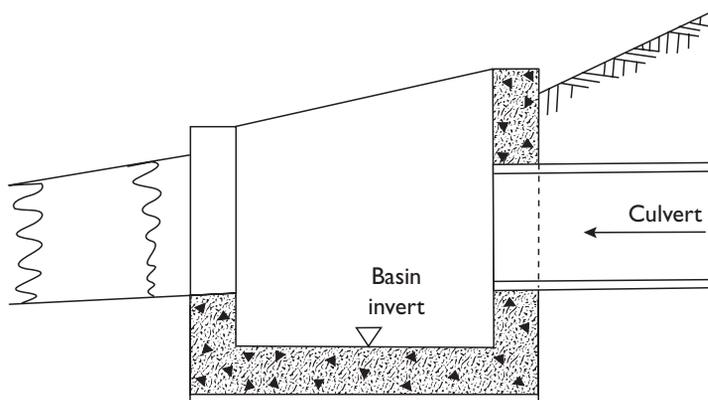


Figure 20.9 Outfall basin.

Source: *International Road Maintenance Handbook* (PIARC 1994).

Replacing
guideposts

causeway structure. Cracks should be cleaned and filled with a bituminous mortar. Erosion cavities upstream or downstream of the drift or causeway should be filled with large stones or riprap. Serious, or recurring erosion may be controlled with gabion mattresses.

Guideposts, marking the water crossing, may be damaged or missing as a result of accidents, flood damage or vandalism. This is dangerous for the traffic when the crossing becomes submerged during floods. Missing or damaged guideposts should be replaced before heavy rains.

20.7.5 Drainage pipes and manholes

Drainage pipes and manholes are mainly used in built-up areas and are often part of the sewerage system. The *International Road Maintenance Handbook* has more details on their maintenance.

20.8 Traffic control devices

Traffic control devices include road signs, guideposts, kilometre-markers, guardrails, and pavement markings. Traffic control devices should be kept in a condition similar to that at the original installation in order to serve their intended function. Maintenance of traffic control devices include

- Routine activities
 - cleaning;
 - repainting;
 - repairing in the workshop.
- Periodic activities
 - replacing guardrails;
 - pavement marking;
 - replacing traffic signs;
 - repairing or relocating kilometre markers.

The *International Road Maintenance Handbook* provides details about their maintenance.

20.9 Implementation

Road maintenance may be undertaken by own staff, contractors, or self-help groups (Beenhakker *et al.* 1987).

Own staff

Traditionally, most maintenance activities have been undertaken by the road administration itself using its own staff. This approach may be used for implementing both equipment-based and labour-based maintenance. The work may be organized as an 'area' system or a 'patrol gang' system. The area system concentrates all the maintenance resources at one location. A maintenance force of approximately 10 staff are responsible for a network of 100–300km of road. This system is most

suitable for a network of roads grouped within a relatively limited area. It is easy to administer because the crew is based at one site only. The patrol gang system includes several units, each consisting of three or more staff and a motor truck. Each patrol inspects and maintains up to 150km of road. A field supervisor directs a number of units and allocates tasks to combined mobile gangs.

Reliance on private contractors for road maintenance is increasing. Evidence shows that contractors, selected through competitive procurement, normally provide labour and equipment at a lower cost and with less delay than government organizations (see Chapter 22). It is obvious that periodic maintenance activities, being rather similar to construction activities, are well suited for contracting. However, the use of contractors for carrying out routine maintenance has proved to be more difficult. In many emerging and developing countries, routine maintenance by labour-based methods has been successfully handled by single contractors, known as 'lengthmen'.

Self-help maintenance is maintenance undertaken voluntarily, without payment, by the local population. It is normally used on rural access roads where the local populations will suffer directly from inadequate road maintenance. Self-help maintenance is most likely to succeed when the road was constructed by self-help. Few self-help maintenance operations exist without some government assistance.

Contractors

Self-help

References

- Beenhakker, H.L., Carapetis, S., Crowther, L. and Hertel, S. (1987) *Rural Transport Services: A Guide to their Planning and Implementation*, London: Intermediate Technology Publication.
- Gichaga, F.J. and Parker, N.A. (1988) *Essentials of Road Engineering with Reference to Warm Climates*, London: Macmillan.
- Heath, W. and Robinson, R. (1980) 'Review of published research into formation of corrugations on unpaved roads', *Supplementary Report 610*, Crowthorne: Transport and Road Research Laboratory.
- PIARC (1994) *International Road Maintenance Handbook: Practical Guidelines for Rural Road Maintenance*, Crowthorne: Transport Research Laboratory for the Permanent International Association of Road Congresses.
- TRRL Overseas Unit (1985) 'Maintenance techniques for district engineers', *Overseas Road Note 2*, 2nd edn, Crowthorne: Transport and Road research Laboratory.
- TRL and DFID (1999) 'A guide to the pavement evaluation and maintenance of bitumen surfaced roads in tropical and sub-tropical countries', *Overseas Road Note 18*, Crowthorne: Transport Research Laboratory.

The HDM-4 road investment model

Henry Kerali

21.1 Introduction

Road investment models

Road investment models can be used to assist with economic appraisal, the preparation of maintenance and investment programmes, and with strategic planning. Typically, these models estimate road construction, maintenance and user costs for a specified analysis period to enable life cycle costs and benefits to be determined for different sets of assumptions. Models simulate the interaction between pavement construction standards, maintenance standards and the effects of the environment and traffic loading in order to predict the annual trend in road condition. These, together with the geometric standards of the road, have a direct effect on vehicle speeds and on the costs of vehicle operation, travel time and vehicle accidents. The assessment of the socio-economic benefits of road investments, however, is difficult to quantify in monetary terms, so is usually undertaken separately. The complexity of these interactions and the need to evaluate a large number of investment options in an economic appraisal means that the use of computer-based models is essential. The best-known example of a road investment model is the Highway development and management tool (HDM-4) (Kerali 1999; PIARC 1999).

Application

HDM-4 may be used to assist with the selection of appropriate road design and maintenance standards that minimize the total transport cost (see Figure 6.6). Computer-based investment models are also particularly helpful for undertaking sensitivity analysis into the affect on investment decisions of uncertainties in traffic forecasts, works costs, discount rates, etc., and for other analyses where a large number of options need to be examined.

Background to the development of HDM-4

The scope of HDM-4 has been broadened over earlier versions of the model to extend its use beyond traditional project appraisals. It now provides a range of tools for the analysis of road management and investment alternatives. It has been developed by an international consortium, with emphasis placed on collating and applying existing knowledge, rather than undertaking major new empirical studies. HDM-4 extends the scope of the existing models in four key areas (Kerali *et al.* 1998):

- *Project appraisal* for the economic appraisal of road maintenance, rehabilitation, upgrading and new construction through life cycle analysis of proposed road investments.
- *Works programming* for the preparation of multi-year and rolling programmes for road network management and development to facilitate the preparation of medium-term budgets for road network maintenance and extension.

- *Strategic planning* for policy development, long-term resource allocation plans and road network planning.
- *Software environment* by the provision of a user-friendly system, built around a set of modules with the capacity to cope with a wide range of data requirements and user skill levels.

HDM-4 also incorporates improved technical relationships mainly by updating and calibrating existing relationships to best current knowledge in the areas of

Technical relationships

- traffic congestion, including the impacts of slow-moving non-motorized traffic;
- wider range of pavement types and structures, particularly with the incorporation of portland cement concrete;
- cold climate effects on pavements;
- traffic safety;
- environmental effects, in terms of energy consumption and vehicle emissions.

21.2 Types of analysis

21.2.1 Project analysis

Project analysis is concerned with the following:

Purpose

- Evaluation of one or more road projects or investment options.
- The application analyses a road link or section with user-selected treatments, with associated costs and benefits projected annually over the analysis period.
- Economic indicators are determined for the different investment options.

Project analysis may be used to estimate the economic or engineering viability of road investment projects by considering factors such as the following:

Project appraisal

- the structural performance of road pavements;
- life-cycle predictions of road deterioration, road works effects and costs;
- road user costs and benefits;
- economic comparisons of project alternatives.

Typical appraisal projects would include the maintenance and rehabilitation of existing roads, widening or geometric improvement schemes, pavement upgrading and new construction.

21.2.2 Programme analysis

This deals primarily with the prioritization of a defined long list of candidate road projects into a one-year or multi-year work programme under defined budget constraints. The candidate road projects are selected as discrete segments of a road network. The selection criteria will normally depend on the maintenance, improvement or development standards that a road administration may have defined (e.g. from the output produced by the strategy analysis application). Examples

Purpose

of selection criteria that may be used to identify candidate projects include the following:

- Periodic maintenance thresholds; for example, reseal pavement surface when damage reaches 20 per cent.
- Improvement thresholds; for example, widen roads with volume/capacity ratio greater than 0.8.
- Development standards; for example, upgrade gravel roads to sealed pavements when the annual average daily traffic exceeds 200 vehicles per day.

HDM-4 application

When all candidate projects have been identified, HDM-4 can then be used to compare the life cycle costs predicted under the existing regime of pavement management (i.e. the ‘without-project’ case) against the life cycle costs predicted for the periodic maintenance, road improvement or development alternative (i.e. the ‘with-project’ case). This provides the basis for estimating the economic benefits that would be derived by including each candidate project within the programme.

Problem definition

For programme analysis, the aim is to seek that combination of treatment alternatives, across a number of sections in the network, that optimizes an objective function under budget constraint. If, for example, the objective function is to maximize the net present value (NPV), the problem can be defined as

Select that combination of treatment options for sections that maximizes NPV for the whole network, subject to the sum of the treatment costs being less than the budget available.

Programme preparation

The programme analysis application may be used to prepare a multi-year rolling programme, subject to budget constraints. The incremental NPV/cost ratio provides an efficient and robust index for prioritization purposes. The use of this is more appropriate than ranking criteria such as the NPV, internal rate of return (IRR), or predicted pavement condition attributes (e.g. road roughness). NPV/cost satisfies the objective of maximizing economic benefits for each additional unit of expenditure from the available budget invested (Kerale and Mannisto 1999).

21.2.3 Strategy analysis

Concept

The concept of strategic planning of medium to long-term expenditures requires that a road organization should consider the requirements of its entire road network asset. Thus, strategy analysis deals with networks or sub-networks managed by one road organization. Examples of road networks are the main (or trunk) road network, the rural (or feeder) road network, urban (or municipal) road network, etc. Examples of sub-networks are all motorways (or expressways), all paved (or unpaved roads), different road classes, etc.

Section identification

A key difference between strategy analysis and programme analysis is the way in which road links and sections are identified. Programme analysis deals with individual links and sections that are unique physical units identifiable from the road network. In strategy analysis, the road system essentially loses its individual link and section characteristics by grouping all road segments with similar characteristics, and then basing the analysis on these groups.

In order to predict the medium to long-term requirements of a road network or sub-network, HDM-4 applies the concept of a road network matrix. This comprises categories of the road network grouped according to the key attributes that most influence pavement performance and road user costs. Although it is possible to model individual road sections in a strategy analysis, most road administrations will be responsible for several thousand kilometres of roads, thereby making this approach cumbersome. The road network matrix can be defined instead to represent the most important factors affecting transport costs in the country. A typical road network matrix could be categorized according to the following:

Network
matrix

- traffic volume or loading;
- pavement types;
- pavement condition;
- environment or climatic zones;
- functional classification (if required).

For example, a road network matrix could be modelled using the following:

Example

- three traffic categories (high, medium, low);
- two pavement types (asphalt concrete, surface treatments);
- three pavement condition levels (good, fair, poor).

In this case, it is assumed that the environment throughout the study area is similar and that the road administration is responsible for one road class (e.g. main roads). The resulting road network matrix for this would therefore comprise ($3 \times 2 \times 3 =$) 18 representative road sections. There is no limit to the number of representative sections that can be used in a strategy analysis. The trade-off is usually between a simple representative road network matrix, which would provide relatively coarse results, compared with a detailed road network matrix with a large number of representative sections that would increase accuracy.

Strategy analysis may be used to prepare medium to long range planning estimates of expenditure needs for road development and conservation under different budget scenarios. For this, estimates are produced of expenditure requirements for medium-to long-term periods usually of 5–40 years. Typical applications of strategy analysis by road administrations include the following:

Applications

- Medium- to long-term forecasts of funding requirements for specified target road maintenance standards.
- Forecasts of long-term road network performance under varying levels of funding.
- Optimal allocation of funds according to defined budget heads; for example, routine maintenance (recurrent) and development (capital) budgets.
- Optimal allocations of funds to sub-networks; for example, by functional road class (main, feeder and urban roads, etc.), or by administrative region.
- Policy studies such as impact of changes to the axle-load limit, pavement maintenance standards, energy balance analysis, provision of non-motorized transport facilities, sustainable road network size, evaluation of pavement design standards, etc.

Methodology As with programme analysis, the aim is to select investment options that maximize an objective function, typically the NPV.

21.3 Structure of HDM-4

21.3.1 HDM-4 modules

Structure The overall structure of HDM-4 is illustrated in Figure 21.1. The three analysis tools ('project', 'programme' and 'strategy') operate on data defined in one of four data managers:

- *Road network* – defines the physical characteristics of road sections in the network or sub-network to be analysed.
- *Vehicle fleet* – defines the characteristics of the vehicle fleet that operates on the road network to be analysed.
- *Road works* – defines maintenance and improvement standards, together with their unit costs, which will be applied to the different road sections to be analysed.
- *HDM configuration* – defines the default data to be used in the applications; a set of default data is provided when HDM-4 is first installed, but users can modify these to reflect local environments and circumstances.

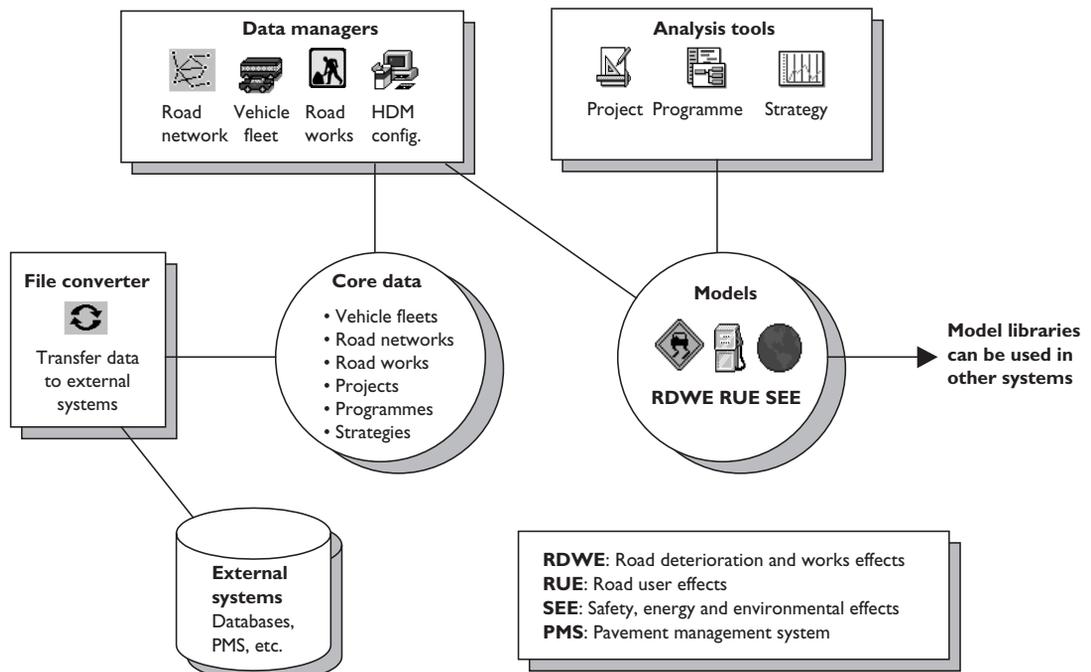


Figure 21.1 HDM-4 system structure.

Source: Kerali (1999).

Technical analysis within the HDM-4 is undertaken using four sets of models: **Models**

- *RDWE* ('Road deterioration and works effects') – predicts pavement deterioration for bituminous, concrete and unsealed roads, and simulates the effects of road works on pavement condition and determines the corresponding costs.
- *RUE* ('Road user effects') – determines costs of vehicle operation, road accidents and travel time.
- *SEE* ('Social and environment effects') – determines the effects of vehicle emissions and energy consumption.

Technical relationships need to be calibrated to match local conditions using country-specific default data, and the 'HDM configuration' module facilitates this.

HDM-4 can interface with external systems such as the following:

External interface

- *Databases* – road network information systems, pavement management systems, etc., through intermediate import/export files.
- *Technical models* – accessed directly by external systems for research applications or other studies.

The modules' import and export functions provide a mechanism for data transfer between existing databases and HDM-4. The data exchange format uses standard data file formats to encourage its wide adaptation by road organizations.

21.3.2 Data requirements

In terms of data requirements in HDM-4, use is made of the concept of 'information quality levels' (IQL) recommended by the World Bank (Paterson and Scullion 1990). Project level analysis data is specified in terms of measured defects (IQL-II), whereas the specification for strategy and programme analyses can be more generic (IQL-III). For example, for project level analysis, road roughness would be specified in terms of its value measured using the 'international roughness index' (m/kmIRI) but, for strategy and programme analyses, roughness could be specified as 'good', 'fair' or 'poor'. The relationship between IQL-II and IQL-III level data is user-defined in HDM-4. Thus, HDM-4 applications can work with a wide range of data types and qualities.

Information quality

Since HDM-4 can be used in a wide range of environments, *HDM configuration* provides the facility to customize system operation to reflect the norms in the environment under study. Default data and calibration coefficients can be defined in a manner that minimizes the amount of data that must be changed for each application of HDM-4. Default values are supplied with HDM-4, but these are all user-definable and facilities are provided to enable these data to be modified.

HDM-4 configuration

The *Road networks* data manager provides the basic facilities for storing characteristics of one or more road sections. It allows users to define different networks and sub-networks, and to define road sections, which are the fundamental unit of analysis. The data entities supported within the road network are as follows:

Road networks

- *Sections* – lengths of road over which physical characteristics are reasonably constant.

- *Links* – comprise one or more sections over which traffic is reasonably constant; this is provided for purposes of compatibility of the network referencing system with existing pavement management systems.
- *Nodes* – intersections which connect links or other points at which there is a significant change in traffic, carriageway characteristics or administrative boundaries.

Data maintenance	Facilities are also available for editing, deleting and maintaining data. The approach to network referencing is designed to handle a wide range of external referencing conventions as might be used by other systems with which HDM-4 may need to interface.
Vehicle fleets	The <i>Vehicle fleet</i> data manager provides facilities for the storage and retrieval of vehicle characteristics required for calculating vehicle speeds, operating costs, travel time costs and other vehicle effects. There is no limit on the numbers or types of vehicles that can be specified. Motorcycles and non-motorized vehicles are also included. Multiple vehicle fleet data sets can be set up for use in different analyses. A range of default data is provided.
Road works	The <i>Road works</i> data manager makes use of the concept of ‘standards’ (see Chapter 2). These are the targets or levels of conditions and response that a road management organization aims to achieve. Standards can be defined in terms either of intervention levels or a frequency of action. Road organizations normally set up different standards related to functional characteristics of the road network. Standards are set to enable specific objectives to be met. The <i>Road works</i> data manager provides facilities to define a list of maintenance and improvement standards that are adopted by road organizations in their network management and development activities. The standards defined can be used in any of the three analysis tools: ‘project’, ‘programme’ or ‘strategy analysis’.
Data requirements	Data required for the HDM-4 analyses can be entered into HDM-4 directly for storage in the internal database. These include data on traffic composition and growth, road maintenance and improvement standards, unit costs and economic analysis parameters (e.g. discount rate, analysis period, etc.). Road network and vehicle fleet data can be set up externally prior to entering into HDM-4. Some data, such as maintenance standards and configuration data can be copied between different HDM-4 applications.
Importing and exporting data	In addition, data required for HDM-4 analyses can be imported from existing data sources. The data import into HDM-4 (as well as the export from HDM-4) is organized according to the data managers, described earlier (i.e. road networks, vehicle fleets, maintenance and improvement standards, HDM configuration). Data transformation rules may need to be implemented for converting the data held in the external database to the format used by HDM-4. For example, pot-hole data in the external database may be recorded in terms of the percentage area of the pavement surface. This would need to be converted to the equivalent number of ‘standard pot-hole units’ (10 litres by volume) required in HDM-4. Similarly, other data required by HDM-4, such as pavement deterioration calibration factors, can be provided as pre-defined default values according to the type of pavement, road class and other defined factors.

21.4 Components of HDM-4

21.4.1 Life cycle analysis

The underlying operation of HDM-4 is similar for each of ‘project’, ‘programme’ or ‘strategy analysis’. In each case, HDM-4 simulates total life cycle conditions and costs for an analysis period under a user-specified scenario of circumstances. The primary cost set for the life cycle analysis includes the costs of capital investment, maintenance and vehicle operating costs, to which travel time costs can be added as an option. The broad concept of the life cycle analysis is illustrated in Figure 21.2, and this is applied to predict the following over the life cycle of a road pavement, which is typically 15–40 years:

Analytical
framework

- Road deterioration
- Road work effects
- Road user effects
- Socio-economic and environmental effects.

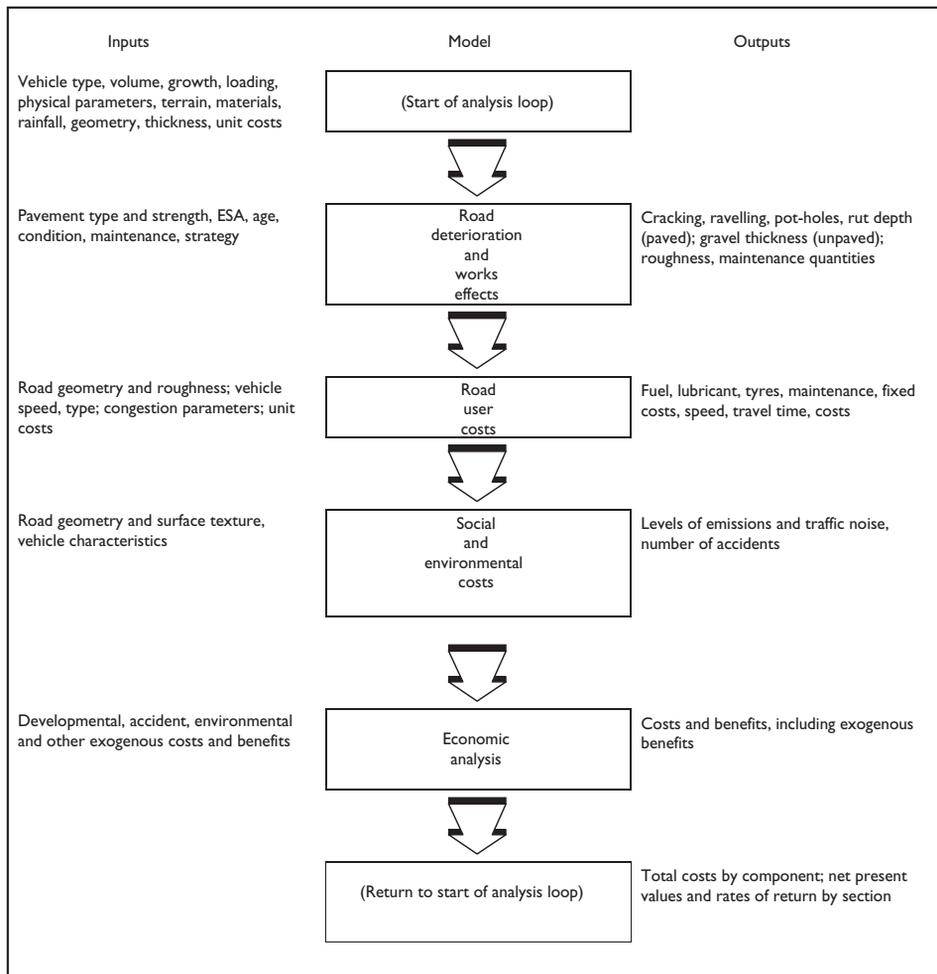


Figure 21.2 Life cycle analysis using HDM-4.

Source: Odoki and Kerali (1999).

Once constructed, road pavements deteriorate as a consequence of several factors. Most notable are as follows:

- Traffic loading
- Environmental weathering
- Effect of inadequate drainage systems.

HDM-4 simulates, for each road section, year-by-year, the road condition and resources used for maintenance, as well as the vehicle speeds and physical resources used as a result of operating vehicles on the road. Interacting sets of costs, related to those incurred by the road administration and those incurred by the road user, are determined for each year and then discounted to give present values. Costs are determined by first predicting physical quantities of resources consumed and then by multiplying these quantities by their unit costs or prices. Economic benefits are then determined by comparing the total cost streams for various maintenance and construction alternatives with a base case ('do-nothing' or 'do-minimum' alternative), usually representing minimal routine maintenance.

Evaluation of alternatives

HDM-4 is designed to make comparative cost estimates and economic analyses of different investment options. It estimates the costs for a large number of alternatives year-by-year for a user-defined analysis period, discounting the future costs. Rates of return, NPVs or first year benefits can also be determined. In order to make these comparisons, detailed specifications of investment programmes, design standards and maintenance alternatives are needed, together with unit costs, projected traffic volumes and environmental conditions.

21.4.2 Road construction cost

HDM-4 does not determine the costs of construction or any other major road improvements. These are specified directly by the user.

21.4.3 Road deterioration

Dependencies

The rate of pavement deterioration is affected directly by the standards of maintenance applied to repair defects on the pavement surface, such as cracking, ravelling and pot-holes. Deterioration is also affected by the treatments that are used to preserve the structural integrity of the pavement, such as surface treatments or overlays, thereby permitting the road to carry traffic in accordance with its design. The overall long-term condition of road pavements depends directly on the maintenance or improvement standards applied to the road. Figure 21.3 illustrates the predicted trend in pavement performance represented by the riding quality, often measured in terms of the international roughness index (m/kmIRI). Maintenance standards define a limit to the level of deterioration that a pavement is permitted to attain. Consequently, in addition to the capital costs of road construction, the total costs incurred by a road administration depends on the standards of maintenance and improvement applied to road networks. Note that the accuracy of the predicted pavement performance depends on calibration applied to adapt the default HDM-4 models to local conditions.

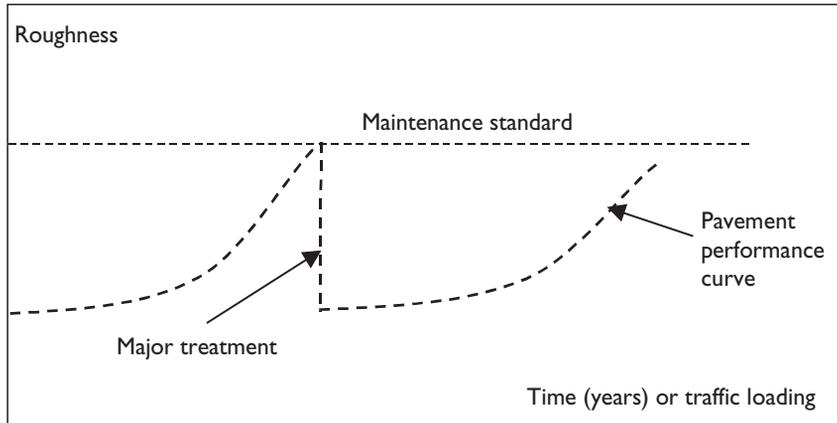


Figure 21.3 The conceptual impact over time of maintenance treatments on pavement condition (roughness) relative to a maintenance standard.

HDM-4 includes relationships for modelling road deterioration and road works effects. These are used for predicting annual road condition and for evaluating road works strategies. As an example, some of the relationships relating to the deterioration of asphalt pavements are described. In this case, road deterioration is predicted through the following eight separate distress modes (Odoki and Kerali 1999):

- Cracking
- Ravelling
- Pot-holing
- Edge break
- Rutting
- Roughness
- Texture depth
- Skid resistance.

The two general classes of model used for road deterioration and works effects analyses are 'mechanistic' and 'empirical'. Mechanistic models are based on fundamental theories of pavement behaviour. They are usually very data-intensive and rely on parameters that are sometimes difficult to quantify in the field. Empirical models are usually based on statistical analyses of locally observed deterioration trends, and may not be applicable outside the specific conditions upon which they are based. To overcome these problems, a 'mechanistic-empirical' approach has been adopted (Paterson 1987). This identified the functional form and primary variables from first principles, and then used statistical techniques to quantify behaviour within this framework. This had the advantage that the resulting models combined both the theoretical and experimental bases of mechanistic models with the behaviour observed in empirical studies.

Relationships

Classes of model

Predictive models	<p>There are two types of models that can be used for predictive purposes:</p> <ul style="list-style-type: none"> ● Absolute models ● Incremental models. <p>Absolute models predict the condition (or distress) at a particular point in time as a function of the independent variables. Incremental models give the change in condition from an initial state as a function of the independent variables. In order to model road deterioration properly, homogeneous road sections need to be identified in terms of physical attributes and condition so that a particular set of relationships can be applied. The basic unit of analysis in HDM-4 is therefore the homogeneous road section, to which several investment options can be assigned for analysis.</p>
Deterioration	<p>The rate of deterioration is a function of the initial pavement design standard, traffic loading, maintenance standards and the effects of the environment. The amount of maintenance carried out in a given year depends on the maintenance standards and the predicted road condition. The annual cost of road works is calculated from the amounts of routine maintenance, periodic maintenance and any road improvements applied in a given year.</p>
Factors	<p>Pavement performance is modelled as a function of several factors. Two of the most important for asphalt-surfaced roads are</p> <ul style="list-style-type: none"> ● pavement strength; ● road roughness.
Pavement strength	<p>Pavement strength is represented by the 'structural number' (SN). The concept of pavement structural numbers is described in more detail in Chapter 15. The structural number of a pavement is defined by an empirical relationship in terms of the thickness and strength of each pavement layer. The pavement layer strength depends mainly on the type and quality of the constituent materials. The structural number of a pavement includes the contribution to pavement strength made by the subgrade. In general, a pavement with high structural number will have a low rate of deterioration under the same regime of traffic and environmental loading.</p>
Roughness	<p>Road roughness is the second most important parameter used in modelling asphalt pavement performance. It is of particular significance in the estimation of vehicle operating costs. It represents the unevenness of a road surface, and is the primary cause of wear and tear on vehicles. The high dependence of road user costs on roughness, and of the change in roughness on the pavement structural number, means that economic analysis results can be significantly affected by the values assigned to these two parameters. The roughness model for asphalt pavements in HDM-4 has several independent variables, including cracking, disintegration, deformation, effects of pavement surface maintenance, structural deterioration of the pavement and the environmental/natural rate of progression. The structural component of roughness relates to the deformation in the pavement materials under the shear stresses imposed by traffic loading. As an example, the incremental change in roughness due to structural deterioration during the analysis year, ΔRI (m/kmIRI), is given</p>

by (Odoki and Kerali 1999):

$$\Delta RI = a_0 \exp(M \cdot K_{gm} AGE3)(1 + SNPK_b)^{-5} YE4$$

where M is the environmental coefficient; K_{gm} a calibration factor for environmental coefficient; AGE3 the pavement age since last overlay (rehabilitation), reconstruction or new construction (years); YE4 the annual number of equivalent standard axles (millions/lane); and $SNPK_b$ the adjusted structural number of pavement due to cracking at the end of the analysis year:

$$SNPK_b = \text{MAX}[(SNP_a - dSNPK), 1.5]$$

Here, SNP_a is the adjusted structural number of pavement at the start of the analysis year and $dSNPK$, the reduction in adjusted structural number of pavement due to cracking:

$$dSNPK = K_{snpk} a_0 \{ \text{MIN}(a_1, ACX_a) HSNEW + \text{MAX}[\text{MIN}(ACX_a - PACX, a_2), 0] HSOLD \}$$

where K_{snpk} is a calibration factor for SNPK; ACX_a the area of indexed cracking at the start of the analysis year (percentage of total carriageway area); HSNEW the thickness of the most recent surfacing (mm); PACX the area of previous indexed cracking in the old surfacing (percentage of total carriageway area); and HSOLD the total thickness of previous underlying surfacing layer (mm).

21.4.4 Road maintenance costs

The different types of road works modelled in HDM-4 include

Dependencies

- routine maintenance;
- periodic maintenance;
- special works;
- improvement works;
- construction works.

The cost of road maintenance is a function of pavement condition, the rate of deterioration and the maintenance standard applied by the road administration. Users can select the types of activities (e.g. pot-hole patching for routine maintenance) as well as the intervention or condition trigger levels and the corresponding unit costs for the activity. The annual cost of maintenance calculated by HDM-4 will therefore depend on the following:

- The values of predicted pavement condition variables.
- Type of intervention specified (scheduled by time, or responsive by pavement condition) that is triggered in each year.
- The unit cost of each operation that is applied.
- The effects of the maintenance operation on pavement condition.

- Example** A typical maintenance standard specified in HDM-4 could comprise the following:
- Routine
 - patch all pot-holes;
 - seal wide cracks covering more than 5 per cent of the carriageway area;
 - annual drainage;
 - miscellaneous maintenance (e.g. grass cutting, line painting, etc.).
 - Periodic
 - reseal every seven years;
 - apply 50 mm overlay when roughness exceeds 5 m/kmIRI.

These maintenance activities would be applied in years when the specified condition triggers are exceeded or in the scheduled years. The total cost of maintenance in any given year is therefore derived from the quantities and unit costs of the maintenance activities that are applied in the year.

21.4.5 Road user costs

Cost components The impacts of the road condition and road design standards on road users can be considered under the following headings.

- Vehicle operation costs (VOC): fuel, tyres, oil, spare parts, vehicle depreciation, utilization, etc.
- Costs of travel time – for both passengers and cargo.
- Costs to the economy of road accidents (loss of life, injury to road users, damage to vehicles and other roadside objects).

VOCs are calculated from the sum of the vehicle resource components, including

- fuel and lubricating oil consumption;
- tyres and spare parts;
- vehicle maintenance labour costs;
- vehicle crew wages;
- vehicle depreciation and interest on capital.

VOC relationships Separate sets of equations are used for the different vehicle types used in HDM-4. For each vehicle type, the models predict average travel speeds as a function of road geometry and road condition. The above VOC components, with the exception of vehicle depreciation and interest, depend largely on road roughness and the geometric characteristics of the road. The consumption of the VOC components are predicted in resource terms. For example, equations for fuel consumption calculate the quantity of fuel consumed over the travel distance. Unit costs for the various resources are specified by the user in order to calculate the annual total costs of vehicle operation. Vehicle depreciation is considered to be a function of the predicted travel time and of the level of vehicle utilization. As an example, the parts

consumption model that is used for each vehicle type k , and for each traffic flow period p , is calculated as follows:

$$PC_{kp} = K0_{pc}[CKM^{kp} (a_0 + a_1 \cdot RI_{adj}) + K1_{pc}][1 + CPCON_k \cdot dFUEL_{kp}]$$

where PC_{kp} is the parts consumption per 1,000 vehicle-km, expressed as a fraction of the average new (or replacement) vehicle price; $K0_{pc}$ the parts consumption rotational calibration factor (default value = 1); CKM^{kp} the average cumulative number of kilometres driven per vehicle type; kp the age exponent in parts consumption model; a_0 the constant term model parameter; a_1 the roughness-dependent model parameter; RI_{adj} the adjusted road roughness (m/kmIRI); $K1_{pc}$ the parts consumption translational calibration factor (default value = 0); $CPCON_k$ the incremental change factor in parts consumption due to vehicle speed-change cycles effects (default = 0.10); and $dFUEL_{kp}$ an additional fuel consumption factor due to vehicle speed-change cycles (i.e. accelerations and decelerations).

Figure 21.4 illustrates the impact of road condition (represented in terms of roughness) on the operating costs of different types of vehicles.

Travel time costs are calculated from average vehicle speeds, travel distances and the unit costs per hour of road users' time. The average vehicle speeds are a function of road roughness, road width and the vertical and horizontal alignment of the road.

Social and environmental effects comprise vehicle emissions, energy consumption, traffic noise and other welfare benefits to the population served by the roads. Although the social and environmental effects are often difficult to quantify in monetary terms,

Time costs

Social and environmental effects

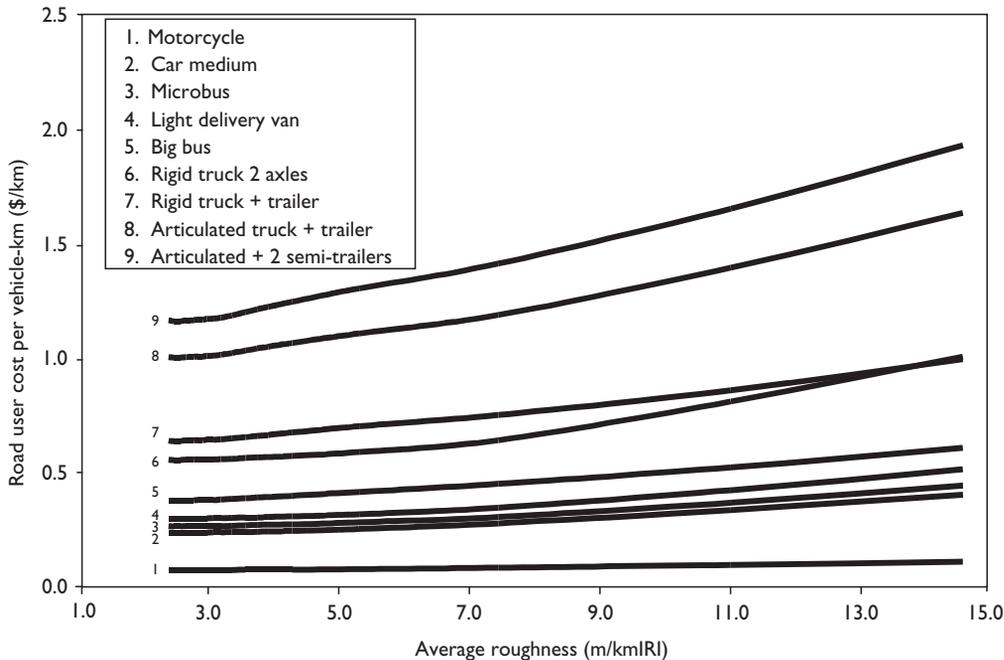


Figure 21.4 Effect of road condition on vehicle operating costs for rolling terrain.

Source: Odoki and Kerali (1999).

they can be specified directly by the user for incorporation within the HDM-4 economic analyses. For example, the quantity of the hydrocarbon emission (g/vehicle-km) is predicted using the following relationship for each motorized vehicle type:

$$E_{-HC} = \frac{3.6 \cdot K0_{ehc}(a_0 + a_1 \cdot K1_{ehc} \cdot IFC)(1 + 0.5a_2 \cdot LIFE) \cdot 10^3}{SPEED}$$

where $K0_{ehc}$ is a calibration factor (default = 1); a_0 , a_1 , a_2 are model parameters depending on vehicle type; $K1_{ehc}$ a calibration factor (default = 1); IFC the instantaneous fuel consumption (millilitres/sec); LIFE the vehicle service life (years); and SPEED the vehicle speed (km/h). Relationships are also included for the cost of traffic accidents.

Non-motorized traffic

In HDM-4, road user effects can be determined for both motorized transport (motorcycles, cars, buses, trucks, etc.) and non-motorized transport (bicycles, human-powered tricycles, animal-pulled carts, etc.).

21.4.6 User-specified costs and benefits

Other costs or benefits directly associated with a road investment may be included in the economic analysis. These usually include benefits accruing from socio-economic developments, such as increased agricultural productivity, industrial output, accessibility benefits, etc. Costs could include the provision of diversion routes, noise barriers and other infrastructure. Such costs or benefits are not calculated by HDM-4 and, therefore, need to be specified directly by the user.

21.4.7 Prioritization

Inadequate budget

There are often situations when the budget available will not be sufficient to undertake all investments shown to have a positive return (i.e. investments with a positive NPV). In such situations, a formal method of selecting investments to be included within the budget can be applied. Prioritization can be applied to a group of investments that meet either of the following conditions:

- Projects that are independent of each other (e.g. road projects from different parts of the country).
- Mutually exclusive projects (i.e. projects that are alternatives to each other) when only one alternative can be selected.

The approach to prioritization can be summarized as follows:

- When sufficient funds are available to undertake all projects
 - select all independent projects with $NPV > 0$;
 - select mutually exclusive project alternatives with the highest NPV.
- When prioritization is to be applied due to shortage of funds
 - select independent projects with the highest NPV/cost ratio;
 - select mutually exclusive projects using the incremental NPV/cost method described below.

The NPV/cost ratio is used as the basis of selecting the best mutually exclusive projects because it measures economic efficiency. The value of the ratio is increased either through a higher NPV, or through a lower cost. 'Incremental analysis' is used to compare the ratio of the increase in NPV to the increase in costs between alternative mutually exclusive projects to a specified marginal ratio. The formula is

Incremental
analysis

$$IBCR = \frac{(NPV_2 - NPV_1)}{(C_2 - C_1)}$$

where IBCR is the incremental benefit/cost ratio; NPV_i the net present value of a project alternative; and C_i the investment costs of a project alternative.

If the above ratio is greater than a specified marginal value, then the project alternative is included among those to be funded. This approach can be used in programme or strategy analysis.

21.5 Examples of applications

Examples of the application of HDM-4 to project, programme and strategy analysis are shown, respectively, in Boxes 21.1, 21.2 and 21.3.

Box 21.1 Example application of HDM-4 in project analysis

This example presents the economic analysis of a project to upgrade an existing engineered gravel road to a paved standard. The existing road traverses rolling terrain, occasionally crossing swampy areas. The current condition is fairly good with an average gravel thickness of 100mm. Consulting services are required to carry out the economic and feasibility study together with preliminary engineering designs.

The proposed project is to pave the road with some improvements in the alignment that will result in a reduction of the road length by about 1.3km. Construction will commence in 2002 and last for two years. Preliminary engineering investigations indicate that the road will have to be designed in two sections. The current traffic on the road sections was observed to be 220 and 242 vehicles per day, on the first and second sections, with a predicted growth rate of 6 per cent per annum up to 2010 and, thereafter, 3 per cent per annum. The key characteristics of the two road sections after construction in 2005 are as given here:

- Section 1:* length 41.8km, width 6.5m, double-surface dressing over lime-stabilized base, AADT of 269;
- Section 2:* length 45.3km, width 6.5m, double-surface dressing over granular base, AADT of 283.

The road administration normally applies annual spot gravel repairs on the existing road, with grading twice a year, and gravel resurfacing at 5-year intervals (the 'without-project' alternative). After paving, the maintenance standard will comprise pot-hole patching and crack sealing (where required), with resealing at seven-year intervals. The economic feasibility of the project

is assessed by comparing this against the ‘without-project’ alternative. The cost of paving is approximately US\$125,000/km (economic – excluding taxes; see Chapter 7) and US\$150,000/km (financial – i.e. including taxes).

The results of the economic analysis at 12 per cent discount rate are summarized in the given table. This shows both the base case scenario (i.e. comparing the ‘with’ and ‘without-project’ alternatives) and the results of a sensitivity analysis with construction costs increased by 25 per cent.

Section	Base case scenario		Increased construction costs	
	NPV (\$)	IRR (%)	NPV (\$)	IRR (%)
Section 1	60,000	12.2	–770,000	9.7
Section 2	750,000	14.4	–150,000	11.6
Overall	810,000	13.4	–920,000	10.7

Sensitivity analysis conducted on these results show that a 25 per cent increase in construction costs would significantly reduce the NPV and IRR. Consequently, the project is at risk of not being feasible if construction costs were to escalate.

Box 21.2 Example application of HDM-4 to programme analysis

This example demonstrates the preparation of a work programme comprising the rehabilitation of a road network in the ‘Western Province’ of a country. The objective is to prepare for funding a prioritized list of investments for sections of road that have been identified to be in poor condition. It is assumed that the list of candidate road sections from Western Province was drawn up following a pavement condition survey of the road network, and the application of pavement condition intervention levels or thresholds to identify those sections in need of some form of rehabilitation. The list would normally contain only those road sections deemed to require some form of periodic maintenance or rehabilitation during the next budget period. Consequently, not all road sections in the network have been selected for the analysis. The AADT observed on each of the road sections is specified together with details of the road sections.

For this example, two maintenance standards have been defined for each road section:

- *Routine maintenance* or ‘do-minimum’ standard
 - crack sealing, pot-hole patching and drainage maintenance, with reconstruction when the pavement has completely failed

- *Rehabilitation* standard

- the above *routine maintenance* standard plus resealing, overlays and reconstruction at specified intervention levels as summarized in the table here

<i>Standard</i>	<i>Works activity</i>	<i>Intervention criteria</i>
Routine maintenance	Patching	Pot-holes ≥ 10 per km
	Crack sealing	Wide cracks $\geq 5\%$
	Drainage maintenance	Scheduled annually
	Reconstruction	Pavement reconstruction at roughness ≥ 12 m/kmIRI
Rehabilitation	Single surface dressing	Total damaged area $\geq 20\%$ of pavement surface area
	Overlay dense graded asphalt	Roughness ≥ 5 m/kmIRI
	Pavement reconstruction	Roughness ≥ 6 m/kmIRI (for asphaltic concrete pavements with high traffic)
	Pavement reconstruction	Roughness ≥ 8 m/kmIRI (for surface dressed pavements with low/medium traffic)

In order to study the effects of budget variations over a three-year period, the *rehabilitation* standard is assigned separately to start at different years of the analysis period. This results in the following four alternatives:

- 1 *Routine maintenance* (base case) from Year 1.
- 2 *Routine maintenance* from Year 1 followed by *rehabilitation* standard assigned from Year 2.
- 3 *Routine maintenance* from Year 1 followed by *rehabilitation* standard assigned from Year 3.
- 4 *Routine maintenance* from Year 1 followed by *rehabilitation* standard assigned from Year 4.

These above assignments of maintenance standards will ensure, for example, that if a road section is not included in the programme for Year 2, it will be a candidate for consideration in Years 3 and 4, whichever gives the highest economic return. The HDM-4 programming methodology aims to select those road sections that should be included within a specified annual budget to maximize the economic benefits.

The results of the analyses indicate that, if there are no budget constraints, the total budget required for rehabilitation in the Western Province would be approximately US\$34 million over a 3-year period. In this case, most of the roads would receive 60mm overlays (see Figure 21.5). However, running HDM-4 with a budget constraint of US\$24 million for rehabilitation over this 3-year period indicates that work on some of the road sections will need to be delayed and others can only receive resealing holding treatments, if the economic return is to be maximized for the available budget (see Figure 21.6). Note that not all sections can be treated in this situation.

HDM - 4

HIGHWAY DEVELOPMENT & MANAGEMENT

Work Programme Unconstrained by Year

Study Name: 1. Western Province (Life Cycle Analysis)

Run Date: 06-11-2002

All costs are expressed in: US Dollar (millions)

Year	Section	Road Class	Length (km)	AADT	Surface Class	Work Description	NPV/CAP	Financial Costs	Cum. Costs
2003	MTS 549 km 0 - 1.0	Primary	19.00	2785	Bituminous	Overlay 60mm at 5 IRI	13.80	0.53	0.53
	MSC 112 km 35 - 80	Primary	21.00	1955	Bituminous	Overlay 60mm at 5 IRI	11.45	0.59	1.12
	MSC 141 km 28 - 49.5	Primary	22.00	2379	Bituminous	Overlay 60mm at 5 IRI	10.93	0.62	1.74
	MAN 203 km 254 - 262.8	Primary	29.00	1968	Bituminous	Overlay 60mm at 5 IRI	10.20	0.81	2.55
	MSC 131 km 16.5 - 26	Primary	10.00	2379	Bituminous	Overlay 60mm at 5 IRI	9.64	0.28	2.83
	MTS 754 km 0 - 1.8	Primary	26.00	1186	Bituminous	Overlay 60mm at 5 IRI	7.27	0.73	3.56
	MSW 959 km 0.1 - 5, 10 - 15.4	Primary	10.00	1398	Bituminous	Overlay 60mm at 5 IRI	7.17	0.28	3.84
	MSE 203 km 80 - 90	Primary	10.00	1316	Bituminous	Overlay 60mm at 5 IRI	6.09	0.28	4.12
	MSW 956 km 10 - 75	Primary	30.00	1830	Bituminous	Overlay 60mm at 5 IRI	5.90	0.84	4.96
	MAN 243 km 0 - 21	Primary	21.00	1153	Bituminous	Overlay 60mm at 5 IRI	5.85	0.59	5.54
	MSC 138 km 5 - 10 & 65 - 70	Primary	10.00	1452	Bituminous	Overlay 60mm at 5 IRI	5.36	0.28	5.82
	MTN 836 km 0 - 13	Primary	13.00	1024	Bituminous	Overlay 60mm at 5 IRI	5.31	0.36	6.19
	MSE 932 km 38.2 - 52.2	Primary	16.00	924	Bituminous	Overlay 60mm at 5 IRI	4.88	0.45	6.64
	MTN 805 km 145 - 152.6	Primary	28.00	964	Bituminous	Overlay 60mm at 5 IRI	4.60	0.78	7.42
2004	MSV 537 km 43 - 45	Primary	26.00	1029	Bituminous	Overlay 60mm at 5 IRI	4.25	0.73	8.15
	MSC 155 km 50 - 55	Primary	45.00	1429	Bituminous	Overlay 60mm at 5 IRI	4.18	1.26	9.41
	MAN 446 km 53.7 - 57.0	Primary	32.00	1123	Bituminous	Overlay 60mm at 5 IRI	3.86	0.90	10.30
	MTN 830 km 8.6 - 20	Primary	11.00	776	Bituminous	Overlay 60mm at 5 IRI	3.60	0.31	10.61
	MSE 334 km 3.8 - 20	Primary	49.00	1065	Bituminous	Overlay 60mm at 5 IRI	2.99	1.37	11.98
	MSW 938 km 3.2 - 10	Primary	47.00	964	Bituminous	Overlay 60mm at 5 IRI	2.74	1.32	13.30
	MAN 214 km 85 - 97.7	Primary	18.00	801	Bituminous	Overlay 60mm at 5 IRI	2.19	0.50	13.80
	MSW 150 km 0.0 - 16.5	Primary	17.00	845	Bituminous	Overlay 60mm at 5 IRI	2.06	0.48	14.28
	MSW 131 km 10 - 16.5	Primary	17.00	489	Bituminous	Overlay 60mm at 5 IRI	1.97	0.48	14.76
	MSW 905 km 15.6 - 170	Primary	42.00	819	Bituminous	Overlay 60mm at 5 IRI	1.75	1.18	15.93
	MSW 901 km 180 - 330	Primary	60.00	1845	Bituminous	Overlay 60mm at 5 IRI	5.06	1.68	17.61
	MAN 409 km 110 - 115	Primary	17.00	560	Bituminous	Overlay 60mm at 5 IRI	1.19	0.48	18.09
	MSW 183 km 10 - 13	Primary	73.00	451	Bituminous	Overlay 60mm at 5 IRI	1.18	2.04	20.13
	MSW 835 km 1.1 - 30	Primary	14.00	605	Bituminous	Overlay 60mm at 5 IRI	0.87	0.39	20.52
MAN 409 km 105 - 110	Primary	59.00	451	Bituminous	Overlay 60mm at 5 IRI	0.81	1.65	22.18	
2005	MAN 203 km 185 - 190	Primary	55.00	1230	Bituminous	Reconstruct at 8 IRI	0.49	9.29	31.47
	MSC 142 km 12.5 - 20	Primary	28.00	1527	Bituminous	Reseal at 20% surface dam	2.11	0.77	32.24
	MSV 449 km 35 - 59.6	Primary	15.00	1153	Bituminous	Reseal at 20% surface dam	0.83	0.39	32.63
	MTN 831 km 0 - 3.7	Primary	14.00	252	Bituminous	Reseal at 20% surface dam	0.29	0.21	32.84
	MTN 748 km 30 - 60	Primary	30.00	721	Bituminous	Reseal at 20% surface dam	0.03	0.39	33.23
								0.42	33.65

HDM-4 Version 1.30

Page -1 of 1

Figure 2.1.5 HDM-4 output for programme analysis when budget is unconstrained.

Notation: 'AADT' = annual average daily traffic; 'NPV/CAP' = ratio of net present value to capital (investment) cost; 'Cum. Costs' = cumulative costs to date.

HDM - 4

HIGHWAY DEVELOPMENT & MANAGEMENT

Work Programme Optimised by Year

Study Name: 1. Western Province (Life Cycle Analysis)
Run Date: 06-11-2002

All costs are expressed in: US Dollar (millions)

Year	Section	Road Class	Length (km)	AADT	Surface Class	Work Description	NPV/CAP	Financial Costs	Cum. Costs
2003	MTS 549 km 0 - 1.0	Primary	19.00	2785	Bituminous	Overlay 60mm at 5 IRI	13.80	0.53	0.53
	MSC 112 km 35 - 80	Primary	21.00	1955	Bituminous	Overlay 60mm at 5 IRI	11.45	0.59	1.12
	MSC 141 km 28 - 49.5	Primary	22.00	2379	Bituminous	Overlay 60mm at 5 IRI	10.93	0.62	1.74
	MAN 203 km 254 - 262.8	Primary	29.00	1968	Bituminous	Overlay 60mm at 5 IRI	10.20	0.81	2.55
	MSC 131 km 16.5 - 26	Primary	10.00	2379	Bituminous	Overlay 60mm at 5 IRI	9.64	0.28	2.83
	MTS 754 km 0 - 1.8	Primary	26.00	1186	Bituminous	Overlay 60mm at 5 IRI	7.27	0.73	3.56
	MSW 959 km 0.1 - 5.1, 10 - 15.4	Primary	10.00	1398	Bituminous	Overlay 60mm at 5 IRI	7.17	0.28	3.84
	MSE 203 km 80 - 90	Primary	10.00	1316	Bituminous	Overlay 60mm at 5 IRI	6.09	0.28	4.12
	MSW 956 km 10 - 75	Primary	30.00	1830	Bituminous	Overlay 60mm at 5 IRI	5.90	0.84	4.96
	MAN 243 km 0 - 21	Primary	21.00	1153	Bituminous	Overlay 60mm at 5 IRI	5.85	0.59	5.54
	MSC 138 km 5 - 10 & 65 - 70	Primary	10.00	1452	Bituminous	Overlay 60mm at 5 IRI	5.36	0.28	5.82
	MTN 836 km 0 - 13	Primary	13.00	1024	Bituminous	Overlay 60mm at 5 IRI	5.31	0.36	6.19
	MSE 932 km 38.2 - 52.2	Primary	16.00	924	Bituminous	Overlay 60mm at 5 IRI	4.88	0.45	6.64
	MSV 537 km 43 - 45	Primary	26.00	1029	Bituminous	Overlay 60mm at 5 IRI	4.25	0.73	7.36
2004	MAN 214 km 85 - 97.7	Primary	18.00	801	Bituminous	Overlay 60mm at 5 IRI	2.19	0.50	7.87
	MSW 901 km 180 - 330	Primary	60.00	1845	Bituminous	Overlay 60mm at 5 IRI	5.06	1.68	9.55
	MTN 805 km 145 - 152.6	Primary	28.00	998	Bituminous	Overlay 60mm at 5 IRI	4.86	0.78	10.33
	MSC 155 km 50 - 55	Primary	45.00	1479	Bituminous	Overlay 60mm at 5 IRI	4.39	1.26	11.59
	MAN 446 km 53.7 - 57.0	Primary	32.00	1163	Bituminous	Overlay 60mm at 5 IRI	4.10	0.90	12.49
	MTN 830 km 8.6 - 20	Primary	11.00	804	Bituminous	Overlay 60mm at 5 IRI	3.81	0.31	12.80
	MSE 334 km 3.8 - 20	Primary	49.00	1103	Bituminous	Overlay 60mm at 5 IRI	3.19	1.37	14.17
	MSW 938 km 3.2 - 10	Primary	47.00	998	Bituminous	Overlay 60mm at 5 IRI	2.96	1.32	15.48
	MSV 448 km 35 - 59.6	Primary	15.00	794	Bituminous	Reseal at 20% surface dam	0.27	0.21	15.69
	MSW 150 km 0.0 - 16.5	Primary	17.00	906	Bituminous	Overlay 60mm at 5 IRI	2.35	0.48	16.17
2005	MSW 131 km 10 - 16.5	Primary	17.00	524	Bituminous	Overlay 60mm at 5 IRI	2.22	0.48	16.65
	MAN 203 km 185 - 190	Primary	55.00	1527	Bituminous	Reseal at 20% surface dam	2.11	0.77	17.42
	MSW 905 km 15.6 - 170	Primary	42.00	878	Bituminous	Overlay 60mm at 5 IRI	2.04	1.18	18.59
	MAN 409 km 110 - 115	Primary	17.00	580	Bituminous	Overlay 60mm at 5 IRI	1.28	0.48	19.07
	MSW 183 km 10 - 13	Primary	73.00	467	Bituminous	Overlay 60mm at 5 IRI	1.27	2.04	21.11
	MSW 935 km 1.1 - 30	Primary	14.00	626	Bituminous	Overlay 60mm at 5 IRI	0.95	0.39	21.50
	MSV 457 km 20 - 28.9	Primary	59.00	467	Bituminous	Overlay 60mm at 5 IRI	0.89	1.65	23.16
	MSC 142 km 12.5 - 20	Primary	28.00	1153	Bituminous	Reseal at 20% surface dam	0.83	0.39	23.55

HDM-4 Version 1.30

Page-1 of 1

Figure 2.1.6 HDM-4 output for programme analysis with optimized work programme when budget is constrained. Notation: As in Figure 2.1.5.

Box 21.3 Example application of HDM-4 to strategy analysis

This example demonstrates the use of HDM-4 for budget forecasting. The objective is to determine the required funding levels for user-defined network performance standards, and to demonstrate the effect of budgetary constraints on the long-term road network performance trends. This involves defining the road network in terms of representative sections and assigning alternative maintenance standards (investment alternatives) to each.

In this example, the entire road network of approximately 8,400 km of main roads has been modelled using 12 representative sections. For bituminous (or paved) roads, these are based on traffic volume ('high', 'medium' or 'low') and road condition ('good', 'fair' and 'poor'), giving a total of nine sections. For unsealed (gravel) roads, the network has been classified by traffic volume alone, since condition may change rapidly in any given year. In this example, the gravel road network is modelled using three sections, representing 'high', 'medium' and 'low' traffic.

Traffic level	Condition	Length (km)	AADT	Roughness (m/kmIRI)	Drainage
<i>Paved road sections</i>					
High	Good	235	6,200	2.5	Excellent
High	Fair	392	5,240	4.0	Fair
High	Poor	437	5,180	6.5	Poor
Medium	Good	306	2,500	2.5	Excellent
Medium	Fair	483	2,300	4.4	Good
Medium	Poor	615	2,060	6.5	Fair
Low	Good	410	1,400	3.0	Good
Low	Fair	670	1,150	4.5	Fair
Low	Poor	720	970	6.1	Poor
<i>Unsealed road sections</i>					
High	Fair	964	300	21.0	Fair
Medium	Fair	1,385	175	9.0	Fair
Low	Poor	1,760	75	12	Fair

Paved road sections

For this example, four investment alternatives are considered for each of the bituminous representative sections, ranging from the provision of routine pavement maintenance only (Alternative 1) to an ideal maintenance case (Alternative 4):

- *Alternative 1*: comprises patching and crack sealing, drainage and miscellaneous maintenance; this represents the base case (or 'do-minimum') alternative.
- *Alternative 2*: includes resealing and reconstruction in addition to the routine maintenance activities in Alternative 1; the objective of this alternative is to adopt relatively inexpensive treatments that will maintain the

existing road in a reasonable condition for as long as possible until the eventual need for reconstruction.

- *Alternative 3*: includes overlays and reconstruction in addition to the routine maintenance activities in Alternative 1; this alternative introduces rehabilitation works at a pre-defined roughness level that will maintain a higher serviceability level.
- *Alternative 4*: this alternative combines the benefits of resealing and overlaying within one standard, which should reduce the required frequency of the relatively expensive overlay works.

Unsealed road sections

The maintenance alternatives are as defined below:

- *Alternative 1*: comprises grading and spot regravelling, both condition responsive. This represents the base case ('do-minimum' alternative).
- *Alternative 2*: includes three works items, grading, spot regravelling and regravelling (resurfacing), all condition-responsive; this alternative aims to maintain a reasonable thickness of gravel surfacing and so provide continuous protection to the pavement subgrade.
- *Alternative 3*: includes the same works items as Alternative 2, but the intervention levels are set to provide a higher level of serviceability that should trigger the road works more frequently.
- *Alternative 4*: introduces upgrading to a paved standard for sections with medium traffic, and widening for sections with low traffic; this represents the ideal investment alternative with frequent grading and regravelling.

Results

The results of the HDM-4 analysis with the above standards indicates the following:

	20-year budget requirements (US\$ million)	Average roughness (m/kmIRI)	
		Paved roads	Unpaved roads
Alternative 1 (base)	230	9.2	15.2
Alternative 2	508	5.3	13.6
Alternative 3	754	3.5	9.4
Alternative 4	1,020	2.7	21.3

References

- Kerali, H.R. (1999) 'Overview of HDM-4', *The Highway Development and Management Series, Volume 1*, Paris: PIARC World Road Association.
- Kerali, H.R. and Mannisto, V. (1999) 'Prioritization methods for strategic planning and road work programming', *Transportation Research Record 1655*, Washington, DC: National Academy of Sciences.

- Kerali, H.R., Robinson, R. and Paterson, W.D.O. (1998) 'Role of HDM-4 in highway management', *Fourth International Conference on Managing Pavements, Durban, 17–21 May 1998, Vol. 2 Implementation*, 801–814.
- Odoki, J.B. and Kerali, H.R. (1999) 'HDM-4 analytical framework and model descriptions', *The Highway Development and Management Series, Volume 4*, Paris: PIARC World Road Association.
- Paterson, W.D.O. (1987) 'Road deterioration and maintenance effects: models for planning and management', *The Highway Design and Maintenance Standards Series*, Washington, DC: Johns Hopkins for the World Bank.
- Paterson, W.D.O. and Scullion, T. (1990) 'Information systems for road management: draft guidelines on system design and data issues', *Infrastructure and Urban Development Department Report INU77*, Washington, DC: The World Bank.
- PIARC (1999) 'HDM-4: highway development and management tool', *The Highway Development and Management Series*, Paris: PIARC World Road Association.

Institutional issues



Chinese in class. (Photo: Bent Thagesen)

Institutional development

Richard Robinson

22.1 Introduction

22.1.1 Current situation

Road transport plays a central role in the economic development of most countries. Large numbers of passengers and large tonnages of goods are moved by road. Industry, agriculture and commerce rely on road transport for both inputs and outputs, and their production relies to a large extent on effective and efficient transport operations. In the emerging countries of Central and Eastern Europe alone, the annual requirement for operations and maintenance requirements is of the order of US\$14 billion. This means that roads represent a large sector with large revenue and expenditure requirements: roads are 'big business'. An effective road transport system is essential to the development of national economies. It is therefore important that road networks are also managed effectively and efficiently.

Role of road transport

However, many studies in developing and emerging countries have highlighted consistently the poor performance of the road sub-sector (Heggie and Vickers 1998; Malmberg Calvo 1998). Road networks in many countries are in poor condition. Although roads are inadequately financed in most countries, lack of financial resources alone cannot explain the poor performance. Institutional arrangements, more widely, are deficient. There is a need to strengthen management, and improve the organizational and regulatory arrangements, in addition to improving financial structures in most cases.

Poor performance

Developing and emerging countries face similar problems with road infrastructure, although the extent and degree vary from country to country. Common issues include the following:

Problems

- network funding that is significantly less than required to sustain networks in acceptable condition;
- in many countries, the structure of funding raised from road users for spending on road network does not reflect the costs imposed on the network by the different types of vehicle; in some countries, the funding structure creates economic distortions;
- network management activities are dominated by the public sector, with staff who are poorly remunerated and motivated, sometimes lacking in competence, and with limited understanding of modern network management practices – resulting in lack of effectiveness and efficiency;

- poorly developed private sector and lack of real competition for much of the procurement undertaken, plus problems of corruption;
- lack of appropriate, relevant and up-to-date management information;
- inappropriate accounting practices that lack transparency;
- politically driven resource allocation that does not reflect economic need;
- inappropriate policy and legislative environment to meet the needs of modern network management;
- lack of public consultation or stakeholder involvement;
- network conditions that are deteriorating over time, resulting in high transport costs and reduced business competitiveness;
- high levels of road traffic accidents;
- inadequate attention to environmental issues;
- problems exacerbated in many countries by high levels of motorization growth.

Road networks

Roads are widely dispersed geographically. Different parts of the network vary in age, and are constructed to different design standards with diverse materials. Often the physical characteristics change over lengths as short as a few hundred metres. Different roads in the network carry different levels of traffic, perhaps varying from tens of thousands of vehicles to fewer than ten vehicles per day. The loads imposed on the network by different types of vehicle also vary. Low-traffic roads may be carrying significant volumes of non-motorized transport and pedestrians. The condition of roads changes continuously as a result of traffic, climate and environmental conditions. Unlike most other network-based infrastructure, roads are an ‘open system’: there is little or no control by the managing authority of the traffic that operates on the network. These disparate factors complicate the road management process, and make the institutional reform and development process much more difficult. Institutional arrangements may differ between major and minor roads within a country, and different solutions may be needed for each of these networks.

Maintenance

The major failing of roads institutions has been that they have not maintained existing networks adequately. Maintenance is a management-intensive activity that requires an effective institutional set-up. Countries have embarked on expensive schemes of road rehabilitation, which could have been avoided if workable systems of maintenance had been established. Every dollar that countries fail to spend on maintenance results in an additional three dollars being spent by road users in higher operating costs because of the resulting poor road conditions (Gwilliam and Shalizi 1996). This imposes higher costs on industry and commerce. Ineffective road management means that finance that is available is spent inefficiently. The failure of governments to allocate sufficient funds has resulted in countries entering into a downward spiral of poorer network conditions, resulting in costs of repair increasing geometrically, with higher running costs limiting road users’ ability to pay more for improved operating conditions. Increased revenue mobilization is required, complemented by improved resource allocation.

22.1.2 Basis of an approach

Effectiveness and efficiency

The aim of institutional development relates to increasing the effectiveness and efficiency of the road sector (Robinson *et al.* 1998). *Effectiveness* measures the

capability to define, agree, and meet appropriate operational objectives. *Efficiency* refers to the ability to undertake activities using minimum resources. Institutional development, or ‘sector restructuring’, involves making changes to existing institutional structures in such a manner that both effectiveness and efficiency are increased over time.

There are no internationally agreed models for sector restructuring, but the building blocks, which are necessary for increasing effectiveness and efficiency are relatively well understood are as follows:

- providing adequate, stable and continuous *finance* – through tariff reform;
- strengthening *management* – through commercialization;
- creating *ownership* and establishing *responsibility* – through regulatory and institutional development.

Building
blocks

Institutional development is discussed under each of these three headings in turn.

22.2 Finance

22.2.1 Issues

It is helpful to consider carefully terms relating to expenditure on roads in the context of the traditional approach to government budgeting. Clarification of concepts is given in Box 22.1.

Roads can be separated into ‘major’ and ‘minor’ road groups (see Chapter 7). Major roads carry relatively high levels of traffic and have primarily an economic function. Minor roads carry low traffic volumes, and have principally a social function. The approach to funding will depend on which type of road network is being considered.

Capital and
recurrent
funds
Major and
minor roads

Effective financing for the operation, maintenance and renewal of roads needs to

- provide a secure source of financing, the availability of which is certain and reliable;
- be independent from political interference on spending decisions, which should be based on need, assessed using economic or other pre-defined criteria;
- establish a direct link between revenue contributions and spending on the road network; with prices paid by road users reflecting the level of service provided;
- collect revenue efficiently.

Principles

Pricing and cost-recovery policies should

- use financing instruments that provide the correct market signals to road users;
- ensure that road administrations use resources efficiently;
- constrain the size of the network to that which is affordable;
- generate sufficient revenues to operate, maintain and renew the road network on a sustainable long-term basis.

It is obvious that, unless funds are made consistently available to meet the costs of road network operation, maintenance and renewal, road networks will decay.

Financial
performance

*Box 22.1 Clarification of budgeting concepts**Capital funds*

Capital funds are those used to provide new facilities or to increase the value, or capital, of existing facilities. It is a political judgement whether or not to provide improved or enhanced facilities, and to pay for these from the public purse, although it is normal that investments meet minimum socio-economic criteria. Governments have a choice whether to make capital investments or not; funding is discretionary. Thus, capital funds should be used for

- development works (new construction).

Recurrent funds

Once a capital facility has been provided, it will need to be operated, maintained and renewed throughout its life. The expenditure of capital funds results in a need to provide recurrent funds over an infinite time frame. This should not be a matter of choice, and no political decision should be involved: funding should be mandatory. Recurrent funds should be used for

- operations
- maintenance works
- renewal works.

Discretionary and mandatory funding

The demarcation of funding in this way indicates that decisions relating to the development of the network are discretionary and part of the political process of government; but that decisions relating to operations, maintenance and renewal of roads should not be part of the political process.

However, road financing is characterized by inadequate and irregular flows of funds, sometimes coming from inappropriate sources. The situation is exacerbated because of poor financial management of the funds that are available, with inappropriate resource allocation, weak accounting procedures and lack of financial accountability. There is now widespread agreement that the root of road financing problems is that roads are seen as a 'free resource'. Roads are funded from general taxation, in the same way as education and health care: users pay taxes and then use of facilities is nominally free. The absence of a financial transaction has encouraged a 'social service' mentality to road management in the past, with resulting inefficiencies in financial management.

22.2.2 Cost recovery**Market-based financing**

Some countries are now moving the financing and management of roads towards an approach that has more in keeping with the commercial values of the private sector (Heggie and Vickers 1998), that is, on a 'market' basis. The aim is to recover the cost of road provision and management from users of the road network, in a manner that

reflects different use of the network in terms of occupancy of road space and of damage caused to the structure of the road. Financing roads in this way requires a fundamental rethink about the nature of roads. Few countries consider roads as part of the market, except in the case of toll roads. However, in the European Union, for example, the *White paper* on transport (European Union 2001) is now proposing weight–distance charges, at least for heavy vehicles. This is an attempt to move to a market-based model for road financing.

Why is a market-based model for road financing appropriate? Under a market-based approach, road users pay directly for using the road. This requires that roads are treated as a ‘utility’ in the same way as telecommunications, power or water supply. Utility users pay an ‘access charge’ to cover fixed costs, but then expect to pay in direct proportion to consumption. Customers can choose when and how much to use the utility, and expect to pay accordingly. Introducing the concept of a ‘utility charge’ has the benefit that there is a *price* for road use (Heggie and Vickers 1998). Market signals are given that encourage drivers to choose whether and how to make a journey. A number of developing and emerging countries are now adopting a utility charging approach to road financing.

Utility charging is the principle that underlies ‘road pricing’. Tolls are the simplest form of road pricing. They are highly visible, and give clear market signals because drivers pay directly for road use. However, tolls cannot be levied on all roads, so other funding mechanisms are needed for the majority of the network. In the long term, electronic road pricing may be adopted for road user charging in industrialized countries, and its operation has already been tested in some places. However, technical difficulties exist with electronic methods and, at least in the short term, other forms of road pricing need to be considered in developing and emerging countries.

In some countries, a fuel levy has been adopted, since this reflects road use directly. Filling stations can post the proportion of the fuel price made up by the fuel levy so that the cost of road use can be seen by motorists. Collection of the levy directly from the refinery is efficient and makes evasion difficult. However, a fuel levy is a poor proxy for road damage because it does not capture the additional costs caused by heavy trucks. Weight–distance charging is used in New Zealand and Switzerland, but a simpler mechanism is to adopt a progressive weight-related annual charge, through a licence or a ‘vignette’. The price of this can be set to match the relative pavement damage caused by different vehicle classes. Utility charging, with a combination of a fuel levy and a progressive weight-related vehicle charge, provides a surrogate market mechanism. The level of charge should be based on need, and be set to enable the network to be managed effectively and efficiently.

22.2.3 Road funds

Some governments are unwilling to adopt an explicit policy of cost-recovery for road use. This is due to a mixture of the following:

- Political considerations resulting from the unwillingness of road users to pay the higher prices that are sometimes required by a policy of full cost-recovery – increases in fuel taxes have resulted in political unrest in some countries, and have even caused governments to fall.

Utility
charging

Road pricing

Interim
charging
mechanisms

Lack of cost-
recovery

- The persistent view that roads are a public good and therefore must be freely accessible to all, not just those who are able to pay.
- A political desire by governments to divert revenues, raised from road users, to meet other government requirements, for example, health, education or defence.

Background

However, many countries are moving towards a system of cost-recovery, where the road administration has its own revenue stream, collected directly from road users, and which is managed according to normal business principles. There is an increased tendency to make revenues as independent as possible from the government taxation system. This is due partly to wide dissatisfaction with systems of generic taxation and with government interference in everyday expenditure decisions. Evidence tends to suggest that, where revenues are linked to expenditure and there is also private sector involvement, road standards are significantly better and costs tend to be lower. Often, the revenues and expenditures are managed through a ‘road fund’ (Heggie and Vickers 1998).

Second-generation funds

Road funds were widely used in the past, but these have been largely discredited being no more than mechanisms for earmarking taxes and reducing fiscal flexibility for ministries of finance. However, pioneering work at the World Bank has resulted in a new ‘generation’ of road funds, where roads are funded on a cost-recovery basis, with revenue collected outside of the government budgeting process, and with expenditures managed according to commercial principles. The requirements for a successful road fund are given in Box 22.2.

22.2.4 Public sector financing of major roads

Road funds are a very powerful tool for financing roads within the public sector. However, in cases where these are not used, the following financing principles should be adopted:

- Users should pay for the costs of network operation, maintenance and renewal through *road user charges* (such as tolls, weight–distance charges, vignettes, vehicle licences, transit fees, fuel levies, or road-related taxes);
- Charges should be structured to recover the cost of road use by different classes of vehicle using the network;
- Collection and expenditure of road user charges should be transparent and accountable;
- Levels of expenditure should be increased until enough money is being spent to keep the network in a technically adequate condition – this requires that network needs are quantified;
- Expenditure on operation, maintenance and renewal should be sourced from road user charges, expenditure on rehabilitation (restoring inadequately maintained sections of road to an acceptable condition) and development (new construction) may be from the budget, sourced from general revenues or loans.

22.2.5 Public sector financing of minor roads

Principles

Although the principle of cost recovery still applies to minor roads, the relatively low levels of motorized traffic on these roads means that funds collected from users will

Box 22.2 Requirements for a successful road fund

Independence of fund

The fund should be independent from the general government budget.

Management of fund

The fund should be managed by a strong and independent Board with clear terms of reference.

Contents of the road fund

The road user charge should typically consist of licence fees, a fuel levy, bridge and ferry tolls, and international transit fees. Note the implication that a 'fuel levy' differs from a 'fuel tax': the assumption is that a tax is collected by government through its normal taxation process. A levy is a genuine road user charge, and may be collected by a non-governmental body.

Collecting the revenues

The road user charge should be collected by the Board and deposited directly into the road fund account without having to pass through the accounts of the Customs and Excise Department or the Ministry of Finance; oil companies can deposit the fuel levy directly into the road fund account, and both licence and international transit fees can be collected under contract.

Setting the level of charging

There should be a formal mechanism for varying the road user charge, and charges should be indexed to ensure that they keep pace with inflation; the Board should either have the power to set the level of charges, based on expert advice, or at least to recommend the level to the Ministry of Finance.

Allocation and disbursement of funds

There should be simple and consistent procedures for allocating and disbursing funds between different administrations entitled to draw upon the fund.

Auditing arrangements

The revenues handled by the fund can be extremely large, so it is important to ensure that these sums of money are accounted for properly; independent financial and technical audits should be instigated to make sure that revenues are collected efficiently, with avoidance, evasion and leakage kept to a minimum, disbursed only according to approved expenditure programmes, and that work is carried out according to specification.

Adapted from: Heggie and Vickers (1998).

rarely be sufficient to cover the costs of road management. However, the minor road network has both an economic function and a social function (see earlier, and Chapter 7). Funding can therefore be on a dual basis to reflect this. A portion of the charges paid by road users could be allocated to the social network to reflect the economic function of the roads. This also reflects the mutual dependency between the

two categories of road. Roads cannot exist in isolation, but as part of a network comprising both major and minor roads. The remainder of the funds can be provided from general taxation to provide for the social function of the network. Minor roads are usually within the jurisdiction of local governments. In addition to user fees, the main sources of funds for minor roads are

- local taxes;
- transfers from central government;
- community financing and cost-sharing.

Local taxes and central government transfers

Local tax revenues are seldom sufficient to fund all of the expenditures assigned to local governments. Roads have to compete with all of the other demands placed upon local revenue sources, so the amount of funding available from this source is a matter for local political priority. Transfers from central government need to be stable from year-to-year and based on objective criteria. However, in most developing and emerging countries, the track record of inter-government transfers to meet these criteria is very poor.

Community financing and cost-sharing

Minor roads are for the benefit of local communities, and responsibility for financing often falls on these beneficiaries. However, cost-sharing has several advantages (Malmberg Calvo 1998):

- It is an effective way of gauging demand for roads.
- Constitutes a financial incentive for communities to organize themselves.
- Expands the revenue base and increases the likelihood that the resulting investments will be affordable.
- Because beneficiaries have put up a substantial proportion of the costs, a sense of ownership is more likely to be derived, and a project is more likely to be sustained.
- It also recognizes that relatively few resources are likely to be available from outside the community.

In Finland and Sweden, local residents are encouraged to form themselves into road co-operatives for managing minor roads (Robinson *et al.* 1998). In Ontario Canada, local roads boards are encouraged and individuals are invited to register as owners of roads. In Namibia and South Africa, there are special arrangements under which commercial farmers receive assistance to enable them to maintain minor roads, providing such roads serve more than one landowner (Heggie and Vickers 1998). These examples provide useful models that could be applied more widely.

Adequacy of finance

Most methods of financing local minor roads are likely to involve some form of cost-sharing. Road funds can be a source of revenue, but must be set up carefully to reduce the possibility of being diverted away from the needs of minor roads. However, introduction of a road fund may result in local governments effectively ceding many of their rights of ownership of the network to the organization administering the fund. Local governments may consider this a 'price worth paying' in return for more reliable and adequate funding levels. However, even when funds are available from central sources, many local authorities lack the resources for any meaningful contribution from their own local revenue, and funds will still be inadequate.

22.2.6 Private sector financing of major roads

A number of options exist for attracting private sector financing for major roads. Many of these involve letting a 'concession' for provision and management of a road, or network of roads. The detailed arrangements for collecting revenue differ. Physical tolls provide a direct mechanism to pay for the road, but note that 'tolls' are not the same as 'private finance'. Tolls can be levied on public roads, and this is common in the United States. Alternatively, roads can be developed using private finance, but no physical tolls are collected. The use of 'shadow tolls' and 'availability fees' in the United Kingdom is an example of this (Robinson *et al.* 1998). The amount of toll necessary depends on the cost of the works and on the traffic level. For a toll rate of 5 US cents per kilometre for passenger cars, the following traffic levels provide general guidance on the minimum required, in general, to recover costs:

- 1,500 vehicles/day to recover toll collection costs only;
- 2,500 vehicles/day to recover operation and maintenance costs;
- 6,500 vehicles/day to recover rehabilitation, operation and maintenance costs;
- 15,000 vehicles/day to recover development (new construction), operation and maintenance.

Toll rates for heavy vehicles are typically between two and five times those for passenger cars. It will be seen that, because of the relatively high levels of traffic necessary to recover costs, private finance will only be suitable for few road construction projects in developing and emerging countries.

22.2.7 Development of financing mechanisms

The development of road sector financing generally needs to progress as illustrated in Figure 22.1. There is a need, over time, to improve the structure of financing, perhaps moving from a system of industrial turnover taxes (as was common in countries of the former Soviet Union), through a taxation system based on road user charges, to a system of road pricing that is outside of the government budget. As the economy matures, there is scope for introducing private sector financing on the most heavily trafficked parts of the network. The aim is to have sufficient funds to meet the ongoing needs of network management.

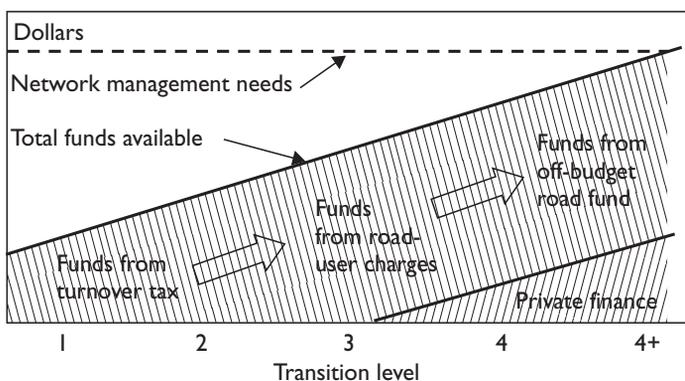


Figure 22.1 Transition of road financing over time.

22.3 Management

22.3.1 Organizational evolution

Road administrations have generally evolved over time, and the World Bank have identified five steps in the evolutionary process, as shown in Box 22.3. The details and steps will differ from country to country depending on the particular socio-political background and the existing situation.

Box 22.3 Evolution of road administration

Stage 1: Traditional construction and maintenance (works) organization

The road organization is centralized and the ministry (of works) manages budgets and projects at a detailed level. The road administration concentrates on technical issues, such as standards and specifications, and works execution. Construction of new roads has priority over maintenance.

Stage 2: Identification of client and supplier functions

The road organization identifies functions of administration and planning (client) and construction and maintenance works (supplier). Traffic safety and axle load control are perceived as problems. The ministry of transport, or equivalent, emerges as a competitor to participate in policy guidance.

Stage 3: Separation of client and supplier functions

The ministry begins to concentrate on policy, and the road administration takes on those client functions that are not assumed by the ministry. There is functional separation, with client functions remaining in the road administration, and the supplier organization reporting either directly to the ministry of transport or to the central management of the road administration. The ministry defines only the mission of the administration, its broad goals, and fixes the budget and pricing rules. A roads board is likely to appear.

Stage 4: Corporatization of the supplier organization

In this stage the supplier organization is divested and, possibly, privatized. A road fund is established to provide for partial autonomy of the road administration. The ministry exercises periodic oversight of the road administration, normally through a board. The road administration is small and manages using modern technology and management systems, and also adopts performance standards.

Stage 5: Corporatization of the client organization

The road administration is corporatized and acquires delegated powers of legal ownership of roads on behalf of government. The administration operates as a private company, subject to oversight from the ministry. Its income source is the road fund paid from road user charges. Management methods treat the road network as a capital asset for which a return on investment must be produced.

Developed from: Talvitie (1997).

22.3.2 Increasing effectiveness and efficiency

The steps in the evolutionary process in Box 22.3 are not random, but have been undertaken with the general aim of increasing the effectiveness and efficiency of organizations, and of the sector more generally. A study, carried out by the World Bank (Israel 1987), investigated the key requirements for achieving management effectiveness and efficiency. The study was undertaken in a large number of organizations in both the public and private sectors. It identified two key requirements:

- increase the ‘specificity’ of an organization;
- subject the organization to competition.

Understanding
the change
process

Specificity is the ability of an organization to identify and focus on its core business, without being side-tracked on to unproductive tasks. Israel (1987) identified several measures for increasing specificity, including the following:

Specificity

- *Objectives* – when well defined, these focus providers of services on increasing efficiency and meeting customer demand; they should be set in terms of output, and defined with as much precision as possible.
- *Time periods* – for meeting objectives can also be defined closely; longer time periods usually imply lower specificity, and a greater likelihood that an activity will be affected by the vagaries of human behaviour or political interference.
- *Procedures* – for achieving objectives can affect specificity; vaguely defined methods, for which there are only general standards, imply that it will be difficult to measure performance and efficiency.
- *Control* – of achievement requires the collection of data so that accomplishment can be verified, and is a result of the ability to specify objectives and methods; controlling achievement is easier with higher specificity activities.

Identification of core business requires the recognition of a customer, or group of customers, and an understanding that the principal aim of the business is to meet customer requirements. This enables business objectives to be set, and management structures and procedures to be put in place that are designed to achieve these objectives. There is also a need to introduce sound business practices and to encourage managerial accountability.

Competitive pressure is also a mechanism for increasing effectiveness and efficiency, since it provides users with choices about how their needs are met, and compels providers to become more efficient and accountable. Competition can be the following:

Competition

- *External* – such as between private contractors in a competitive tender situation.
- *Internal* – competitive pressures can be exerted on an organization by the political establishment, regulatory agencies and by road users, and by managerial measures that create a competitive atmosphere within the organization.
- *Mixed* – where a public sector organization competes with organizations from the private sector – this has been carried out in the road sector in the United Kingdom with great success (Madelin 1994).

Avoidance of monopolistic control	<p>It should be noted that the main issue is the degree of monopolistic control exercised, rather than whether there is public or private sector involvement in the works. Note that a parastatal or private body operating in a monopoly position has little incentive to perform better than a government organization, and both can be much less accountable in terms of prices and level of service. It is, therefore, emphasized that the key to effectiveness and efficiency in this area is <i>competition</i>, not necessarily privatization.</p>
Functional separation	<p>A further mechanism for increasing specificity is to separate organizations, or functions within organizations, into business units. Each of these units has a core business and a defined customer at whom the business is targeted. This can be done at a number of 'levels' in the road sector.</p>
Government and the executive	<p>At sector level, specificity is increased by separating transport from other infrastructure or public works sectors at ministerial level. For example, a Ministry of Transport has higher specificity than a Ministry of Works, because it serves a more focused sector and has a well-defined customer. Within the transport sector, separation of road infrastructure provision and management from transport operations also increases specificity. For example, creation of a road administration separates responsibility for roads from other transport responsibilities of a Ministry of Transport. In this way, there is separation of the government policy and political (ownership) roles from the executive roles of administration, management and works execution.</p>
Client and suppliers	<p>The functions of a road administration can be split conveniently into those for the management of road operations (the 'client' responsibilities for strategic planning, programming, preparation of designs and contracts, and operations management), and those for works execution (the 'supplier' responsibilities). The client role is concerned with specifying activities to be carried out, determining appropriate standards to use, commissioning works, supervising, controlling and monitoring activities. The supplier role is concerned with delivering the defined product to an agreed quality standard, to time and to budget. Arrangements between the parties can be put on a more contractual basis. Such separation clarifies roles, and increases the focus and specificity of action, both of the management and the works execution functions. Both parties have incentives to increase operational performance. Greater benefits in effectiveness and efficiency have been achieved as a result of such separation than from virtually any other type of organizational reform.</p>
<p>22.3.3 Private sector</p>	
Scope for private involvement	<p>An emerging theme of institutional reform is an increasing emphasis on the involvement of the private sector, not only for undertaking physical works on the network, but also in ownership, financing and management. At one end of the spectrum are private sector toll roads, owned and operated by businessmen and entrepreneurs. At the other end, there is a trend to introduce private sector practices and disciplines into road administrations. Many of these are now run using commercial principles, with private sector conditions of service for staff, commercial accounting procedures, and with full management accountability for efficient road operations. All of these measures increase specificity.</p>
Benefits	<p>Use of the private sector for major development and rehabilitation works is widely accepted. However, benefits are apparent for other activities, although</p>

cost-comparisons are difficult because of different accounting practices used in government and in the private sector (Heggie and Vickers 1998; Robinson *et al.* 1998). In those countries, such as New Zealand and the United Kingdom, where the private sector is used for a wide range of activities, typical benefits quoted are as follows:

- Maintenance works: 17 per cent (New Zealand); 10–15 per cent (periodic works in UK); 30–50 per cent (routine works UK);
- Contracted-out management: 20–30 per cent (New Zealand); 15 per cent (UK).

These savings can provide targets for developing and emerging countries, and a World Bank study of contracting-out maintenance (Miquel and Condrón 1991) indicated that many countries are well on the way to achieving similar benefits. Note that the main cause of these benefits is the introduction of competition, rather than simply use of the private sector.

22.3.4 Approach to restructuring

The ability to increase specificity by separating functions and to subject them to competition depends on the particular roles concerned in operating and managing the road sector. Role definitions are given in Box 22.4.

A model of a functionally separated road network management arrangement, with extensive use of competition, is illustrated in Figure 22.2. In this case, each role is performed by a different organization. ‘Client–supplier’ relationships between these organizations are formalized through written agency agreements between public sector bodies, and through contracts where private sector organizations are involved. This approach enables roles to be defined precisely, and performance targets to be set as part of the agreements between clients and suppliers. Thus, there is high specificity. Competition is introduced, wherever possible, between concessionaires, network managers, and contractors. However, national and city road administrations are *de facto* monopolies, so the model assumes that these are state-owned companies, separate from their parent organizations, and operating in a quasi-commercial manner.

Sectoral roles

Model arrangements

Box 22.4 Sectoral roles

- *Owner* – responsible for funding; policy formulation; framing legislation and regulations; etc. (typically a government ministry);
- *Administrator* – responsible for effecting policy by ensuring network performance meets overall political aims of owner (the ‘road administration’);
- *Manager* – specifies, supervises, controls and monitors activities;
- *Contractor* – undertakes works.

Source: Parkman *et al.* (2001).

Note

‘Service provider’ – combines both the manager and contractor roles.

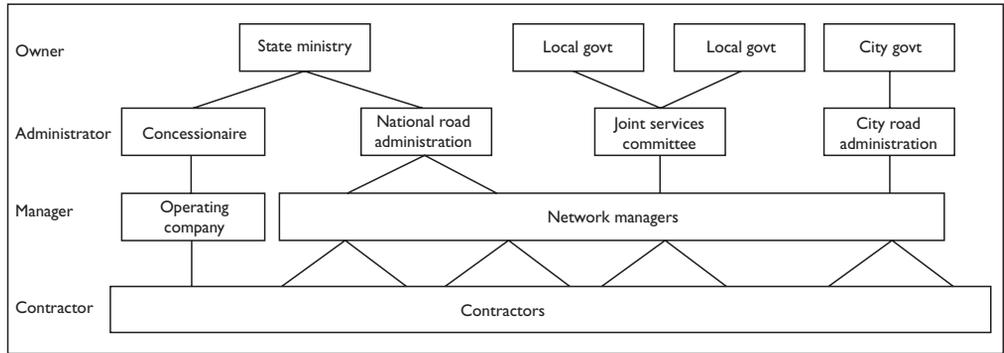


Figure 22.2 Model road network management arrangements.

The privatization of natural monopolies is discouraged. The use of ‘joint services committees’ for administering local government roads recognizes the small size and often limited technical capability of many local government bodies (Malmberg Calvo 1998). It provides a mechanism for obtaining an administration with a minimum critical mass of size and competence.

Organizational transition

Looking back at the Talvitie evolutionary model, illustrated in Box 22.3, it can be seen that as the evolutionary process develops, there is a gradual increase in specificity and competition:

- *Stage 2* (identification of client and supplier functions) – increases *specificity* by separating transport from other public works sectors, and by separation of road provision and management from transport operations.
- *Stage 3* (separation of client and supplier functions) – increases *specificity* by separating road ownership from the administration, management and contracting roles.
- *Stage 4* (corporatization of the supplier organization) – subjects the service delivery/contracting functions to *competition*.
- *Stage 5* (corporatization of the client organization) – subjects the client organization to *competition*.

Figure 22.3 shows how this model is being used to plan the gradual commercialization of road network management in Romania. In Stage 1, the roads organization was part of a monolithic Ministry of Public Works. In 1991, the National Administration of Roads (NAR) was formed, which was functionally separate from its parent Ministry. In 1996, the periodic works units were separated from NAR and corporatized as ‘joint stock companies’. These were companies in which the government owned all shares. They were formed in this way to enable them to build up experience of working in a competitive market, whilst still having some government protection. In 2002, these companies were privatized, and the routine maintenance organizations were corporatized. The figure shows how functional roles are gradually being moved into the private sector. The goal is to achieve the status indicated by ‘4+’, which is similar to the way the road sector functions in New Zealand,

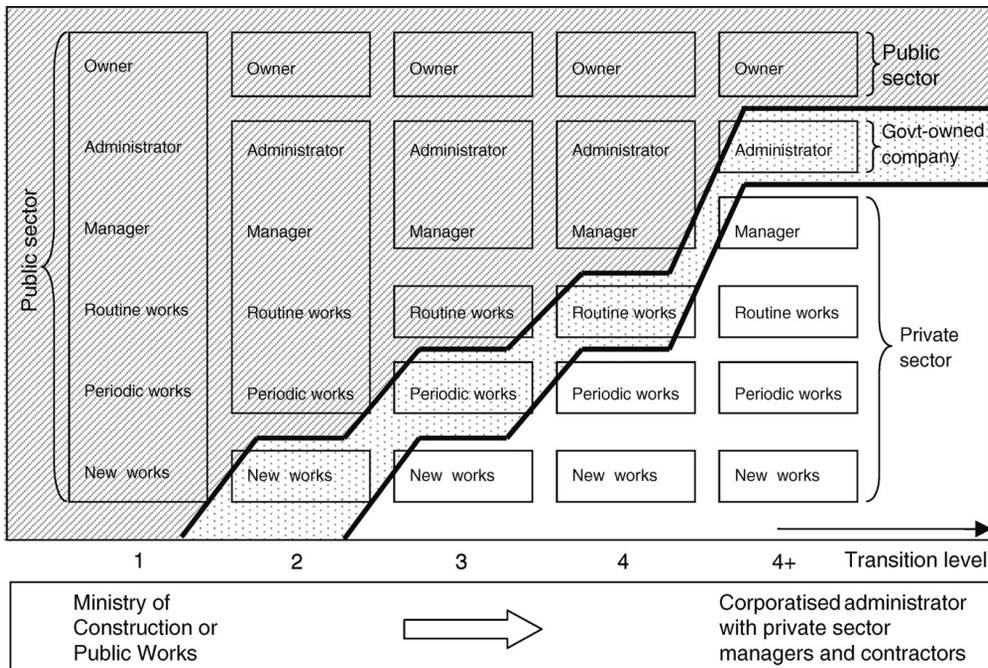


Figure 22.3 Five-stage strategy for restructuring the organization of road network management in Romania.

Sweden and the United Kingdom. Note that the role of administrator remains a government-owned company. This is because this organization is a natural monopoly, so is not appropriate for privatization.

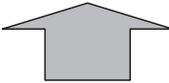
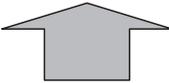
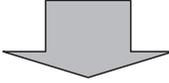
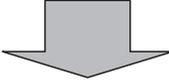
22.3.5 Minor roads

The institutional development of organizations responsible for minor roads can follow the above pattern. However, the process is complicated because many of these organizations are decentralized. Different types of decentralization are possible, reflecting different relationships between the network 'owner' and 'administrator' for funding and management. The best types of decentralization meet the following criteria in terms of capability for financing and managing rural road networks (Malmberg Calvo 1998):

Decentralization

- ability to reflect local priorities in policy formulation and decision making;
- achievement of market discipline (i.e. competition) through their management and procurement arrangements;
- scale of operations in terms of having sufficient critical mass for operations to be effective and efficient;
- simplicity of administration in terms of decision-making chains and other linkages.

Table 22.1 Summary of road network management models for administrative decentralization

Model	Owner	Administrator	Manager	Contractor
Devolved	Local government	Local government		
Bottom-up principal agency	Local government	Central government under an agency agreement	Private sector under contract	Private sector under contract
Devolved and delegated	Local government	Parastatal or private sector under contract		
Joint services committee	Local government	Joint service committee of participating local governments		

Adapted from: Robinson and Stiedl (2001).

Note

'Market decentralization' is where the private sector takes on at least one of the administrator, manager or contractor roles; it is difficult for the private sector to take on the role of owner.

Recommended model

Table 22.1 shows those types of decentralization that have been found to be most effective. In this table, 'delegation' involves the assignment of responsibility through a formal contract; 'agency' involves assignment of responsibility through an agency or framework agreement. Whereas contracts can involve either public or private sector, agency agreements of this nature are normally only between public sector bodies – but there are exceptions. Other types of decentralization are also found in practice, but do not work so well. Note that the joint services committee approach can prove complex because of the involvement of several different authorities; however, if ways can be found of overcoming these difficulties, the approach offers considerable promise, in conjunction with private sector management and contractors.

Funding

This table makes no comment about reliability of funding, which is likely to be similar (and probably poor) under all models. Use of a road fund can overcome this problem, but road funds are very centralized instruments. An important issue for decision makers is whether to adopt a road fund approach to ensure the availability of at least some funding, and to trade this off against the possible resulting loss of authority about how funds are spent.

22.3.6 Management practices

Management practices need to be improved in most road organizations. This can often be facilitated through 'commercialization', which can be considered as the process of transforming the operations of a public sector organization to work in a more business-like manner, and become more effective and efficient. Commercialization does not necessarily involve privatization (see earlier), and the terms should not be confused. The management practices shown in Box 22.5 are typical of those that can be improved, irrespective of the type of organization concerned.

*Box 22.5 Commercialization of management practices**Customers and stakeholders*

Effectiveness and efficiency are increased by gearing services to the stated needs of identified customers. This can be increased further by involving formally all stakeholders in the planning and management functions related to the road network – perhaps through a roads board.

Policy framework

Clear definition of the policy framework increases the effectiveness and efficiency of an organization. This enables business objectives to be set and performance targets to be monitored (see Chapter 2).

Commercial management practices

Improving effectiveness and efficiency by introducing management practices that are similar to those in a commercial organization. These include the need for

- introduction of appropriate management structures and operational procedures;
- strengthening managerial accountability and delegation of responsibility;
- introducing commercial costing and accounting practices, with independent audit;
- addressing human resource requirements, including employing sufficient (and only sufficient) staff with adequate skills;
- introducing transparent and modern procurement practices, with prompt payment of suppliers;
- improving access to information by implementing management systems in the areas of finance and accounting, road management, bridge management, and personnel management;
- adoption of improved methods of resource allocation based on quantified need, aimed at maximizing returns on investment.

Developed from: Heggie and Vickers (1998).

22.4 Ownership and responsibility

Communities affected by road projects have a substantial interest in decisions made in their planning and design, and can often assist with information on local conditions and preferred environmental mitigation measures. Timely consultation and participation are very important, can take many forms, and need to be implemented and institutionalized. There is considerable benefit, more generally, in involving road users and other stakeholders in the road network in the road management process. This may involve the formation of roads ‘boards’. These should have significant private sector involvement, and should participate in, or be responsible for, road network management. The creation of roads boards in several countries has resulted in a change of focus of management activities to gear its services more closely to meet

Public
consultation

- the needs of road users (Heggie and Vickers 1998). This has created a strong pressure to increase the effectiveness and efficiency of road network management arrangements.
- Legislation Policy for institutional reform is implemented mainly through legislation, which may be needed in the following areas:
- Roads Act, defining ownership, roles and responsibility for the road network, and including a classification of the network.
 - Competitive procurement legislation, with provision for enforcement, and covering dispute resolution and arbitration.
 - Concession law allowing provision and operation of public roads by private companies under negotiated commercial contracts.
 - Road fund or utility charging law.
 - Legislation on vehicle emissions and other road transport-related pollution.
 - Fair trading legislation.

22.5 Expected outcomes

Quantifiable benefits Implementing road sector reforms is expected to give rise to a number of benefits. Talvitie (1997), from the World Bank, has noted that the case for restructuring on efficiency grounds is supported by the following general conclusions:

- Functional separation of the administration and management of roads can increase efficiency by 10–15 per cent.
- ‘Optimal’ timing and scheduling of works through a life-cycle approach, as obtained under private sector management, reduces the total road transportation costs (user and administration) by 5–30 per cent.
- Contracting can reduce costs by 5–15 per cent.

Non-quantifiable outcomes With better funding and management, road network conditions should improve over time. This not only brings considerable benefits to road users in terms of convenience and cost savings, but timely road maintenance interventions also reduce future costs for the road administration. Improved network conditions can assist directly in economic development and growth. Sector restructuring can help to create a vibrant private sector, again contributing to economic development. Transfer of responsibilities and staff from the public to private sector, or to government-owned companies, can increase levels of staff remuneration, provide a better working environment, and lead to increased job-satisfaction.

References

- European Union (2001) ‘Fair payment for infrastructure use: a phased approach to a common transport infrastructure charging framework for the EU’, *White Paper*, Brussels: Commission of the European Communities.
- Gwilliam, K. and Shalizi, Z. (1996) ‘Sustainable transport: priorities for policy reform’, *Development in Practice*, Washington, DC: The World Bank.
- Heggie, I.G. and Vickers, P. (1998) ‘Commercial management and financing of roads’, *Technical Paper No. 409*, Washington, DC: The World Bank.

- Israel, A. (1987) *Institutional Development: Incentives to Performance*, Baltimore: Johns Hopkins for the World Bank.
- Madelin, K. (1994) 'Highway maintenance management in Shropshire', *Proceedings of the Institution of Civil Engineers, Transportation*, **105** (May), 97–103.
- Malmberg Calvo, C. (1998) 'Options for managing and financing rural transport infrastructure', *World Bank Technical Paper No. 411*, Washington, DC: The World Bank.
- Miquel, S. and Condron, J. (1991) 'Assessment of road maintenance by contract', *Report INU91*, Washington, DC: The World Bank.
- Parkman, C.C., Madelin, K.B. and Robinson, R. (2001) 'Potential for private-sector road maintenance in developing countries: cases studies', *Transportation Research Record No. 1749*, 1–8.
- Robinson, R., Danielson, U. and Snaith, M. (1998) *Road Maintenance Management: Concepts and Systems*, Basingstoke: Macmillan.
- Robinson, R. and Stiedl, D. (2001) 'Decentralization of road administration: case studies in Africa and Asia', *Public Administration and Development*, **21**, 53–64.
- Talvitie, A. (1997) 'International experience in restructuring road sector', *Transportation Research Record No. 1558*, 99–107.

Training of staff

Richard Robinson and Bent Thagesen

23.1 Purpose of training

Aim

The aim of training is to improve job performance by extending knowledge, developing skills and modifying attitudes, so that individuals can work in the most economical, efficient and satisfying way (Barber 1968). Training is a way of improving performance by changing the way that work is done. It is an indispensable requirement for improving resource allocation and can be used in the areas of

- Management techniques
- Technical matters
- Work methods and practices
- Procurement procedures.

Training is not discretionary; it is an economic necessity. It should also be a continuing function and part of an organization's strategic development.

Scope

At the corporate and the individual level, training may be considered as

- a sequence of human-related activities that enable the road organization to perform delegated responsibilities effectively and efficiently; and
- enabling a staff member, or a group of staff members, to conduct delegated functions and tasks within the road organization competently, and consistent with the corporate performance requirement.

The discussion on training in this chapter is concerned with staff of a road administration, consulting or contracting company, who are already employed or newly recruited. Thus, in this context, training is different from education, and from pre-employment vocational and technical training. The focus on the corporate performance will, to a large extent ensure that an organizational culture is developed, enhancing motivation, co-operation, communication and appropriate management systems.

Strategic role

The strategic role of training is to effect improvements in the way that road organizations' work is actually performed, how equipment is operated and maintained, and how materials are used in the field. Ultimate solutions are long term in nature, but there are things that can and should be done in the short term. There is a general need for training programmes that are properly targeted at those activities that will have

a direct impact and give an immediate return. For many road organizations, these will include increasing productivity and service levels.

There are many reasons why employees need to be trained (Nadler and Nadler 1989), including the following: **Needs**

- there are new employees to the organization;
- to overcome labour shortages;
- to mitigate high turnover of staff;
- to provide new knowledge and skills needed because of
 - changes in organizational practices, including organizational restructuring, decentralization or making more use of the private sector,
 - expansion or diversification of activities; and
- to improve the quality of work, to raise the calibre of staff and to prevent deterioration in performance.

The road organization will normally have a preference for recruitment of technical personnel from nationally recognized universities, colleges and training institutions to obtain the most relevant qualifications and skills possible. There may, however, be a number of constraints. The availability of qualified personnel may be insufficient. The mobility of professionals and technicians may be low. The employment incentives may be inadequate to attract the most qualified personnel. The taught qualifications and skills may not be compatible with the performance requirements of the road organization. There will also be a need to train administrative and clerical staff. In addition, many employees will be at the sub-professional or manual worker level, who will have few, if any, formal qualifications. Training will be needed for these staff, if their work is to be carried out efficiently and effectively. **Recruitment**

In many countries, the curricula at the universities and colleges have not been kept up-to-date, and the education and training have become obsolete. To overcome this, road organizations in some countries have established their own education and training institutions that function at the national level. **Own training institutions**

In countries with a well-developed private sector, there may be a high staff turnover in public organizations as the private sector may offer more attractive employment opportunities. This increases the need for training in the roads organization since staff must be recruited to replace those who leave, and new staff may not possess all the required qualifications and skills. With the increased focus on privatization in recent years, several governments have also organized training programmes targeted at the private sector. The need for training will increase further if the educational background for certain disciplines is limited and the skill base for recruitment is small. **Private sector**

Training is often prescribed as a panacea for transferring knowledge and skills, but it has been stated (Israel 1987) that training has a poor record of achieving institutional development. Undue emphasis has been placed on inputs (such as the number of people trained) rather than impacts (such as the effects of training). Training needs to be formulated with considerable care if sustainable results are to be obtained. In particular, training should always be designed to meet clear objectives, with achievement that is measurable. Training needs to be developed in the context of the organization's wider aims and objectives. **Achievement of training**

23.2 Institutional issues

- Causal factors** There are a variety of factors that may cause ineffective use of staff, other than inadequate training. These include personnel policies resulting from the application of civil service rules, conditions of employment and pay within the public sector organizations, lack of accountability and incentives, the level of the organization's efficiency and structural complexity. These can be compounded as a result of institutional restructuring and privatization. Such factors can affect the general ability to attract, retain, train and motivate technical and managerial staff. The result is often a demoralized work force, with apathy pervading everyone from senior management, through to equipment operators and labourers. Against such a background, human resource performance, and indeed training, is always likely to be ineffective.
- Institutional appraisal** Thus, there is a high degree of correlation between training issues and institutional issues. Experience has shown that an institutional appraisal, covering organizational, managerial and technical issues, is the most effective method of identifying clearly training objectives and institutional constraints (Brooks *et al.* 1989; Robinson 1991). In particular, an institutional appraisal will help to develop a common understanding of what issues training should address so that its objectives will best suit those of the organization. Thus, institutional objectives and arrangements are deciding factors of key importance for the design of training interventions. The training needs assessment and the institutional appraisal should be undertaken together.
- Achievement of objectives** Institutional appraisals can identify the client's objectives in the different areas of its operations. It is then possible to work backwards from these to identify any human resource constraints that might be impeding the attainment of objectives. The intention is to achieve a balance between the organizational objectives and the institutional capacity. This means that the technical and financial resources should correspond to the responsibilities and tasks that the organization is required to carry out.
- Human resource development** The availability of trained engineers, technicians and managers varies from country to country. But even where well-trained staff exist, institutional performance has tended not to improve significantly because these personnel have not been effectively developed, utilized and retained. A further reason why the training record is sometimes disappointing is that training is often seen in isolation from the broader subject of human resource development. Insufficient attention is paid to establishing a framework of priorities to ensure that training is cost-effective and results-orientated. More attention needs to be given to analysis of staff issues before training is planned, and to personnel management after training has been completed. Insufficient analysis of staff resources and needs carries the risk that the wrong type of training may be given, or that people are trained for jobs that are of low priority or do not exist. Poor personnel management means that trained staff are unable to apply what they have learned effectively, and their training is wasted (Relf and Thriscutt 1991). Whereas the earlier focus was on training, it is now apparent that the human resource problem is more a question of the institutional and personnel management issues of utilization, motivation, development and retention (Moeller and Iacono 1990).
- National policy** The above issues should preferably be dealt with at the national level, and be incorporated in policies for the transport and road sectors. Even where a national policy is absent, the process of transition and change may very well already be under

way. There are general trends resulting from the restructuring of government services at large, and in the transport sector in particular that suggest the need for common elements of training. Much more emphasis will need to be given to management training than was the case in the past when technical training tended to dominate. However, there are no standard solutions for institutional development or staff training. Solutions must be based on a pragmatic approach in accordance with the prevailing environment.

23.3 Training types

Training can be delivered internally or externally, or as a combination of both. Internal training is provided in the road organization's own training centre or as in-service training. External training is provided in national or foreign training institutions. Training may be funded locally, or may be project-related and carried out as part of an activity that is funded internationally. There are a number of different types of internal and external training.

Internal and external training

Internal training provides better opportunities than external training for adapting the training interventions to local needs, and it interferes very little with the trainees' social life. Internal training can take place without disrupting the service of the road organization significantly, and it can be conducted in the right context. Internal training may comprise on-the-job training, classroom training and counterpart training.

Internal training

On-the-job training emphasizes practical skills and routines, for example, operating machines, instruments, etc., which are learned by observing and doing the actual work under the supervision of a skilled and experienced operator. A special form of on-the-job training, often used in road construction and maintenance, is the 'training production unit' (TPU). A TPU is attached to a training centre and is delegated the responsibility for construction and maintenance of a certain part of the road network. The TPU is composed of engineers, foremen, equipment operators and mechanics, who undertake the assigned maintenance work. The TPU has all the necessary equipment and tools, and access to workshop facilities. TPUs provide a well-focused form of on-the-job training.

On-the-job training

Classroom training deals with the theoretical aspects associated with particular tasks. On-the-job and classroom training are often related to manual and technical job categories, but classroom work also plays a key role in management training.

Classroom training

Frequently, foreign staff assigned to a development project will have local training counterparts. The counterparts are expected to take over when the foreign staff leave. Counterpart training is an informal way of transferring know-how during the daily co-operation when executing the work. It is normally related to middle and higher management job categories. Unfortunately, the success of counterpart training has been very poor, as discussed in Box 23.1. To be effective (Relf and Thruscott 1991), counterparts should be assigned throughout the project, must be motivated and must have the knowledge and capability to develop to the required level. There must, therefore, be incentives for development, and counterparts should be able to see that they will eventually be engaged in the position for which they are being trained, and will be able to apply the expertise that they have acquired. Best results are likely if, as quickly as possible, the roles of consultant and counterpart are exchanged, with the consultant then filling a largely advisory position.

Counterpart training

*Box 23.1 Problems with counterpart training**Client organizations*

On the part of government, counterparts are often assigned late, or only temporarily, to consultants. Often they are not assigned at all. Sometimes the counterparts may not be suitable for the type of training given or for the post that they may later hold. In some cases, foreign staff simply act as 'gap fillers'. There are often good reasons for the lack of counterparts as this extract from a consultant's report illustrates:

The availability of staff for re-training and upgrading was limited by operational needs and turnover. There was a higher percentage turnover of engineers, technicians and skilled operatives. In most cases, experience has been replaced by inexperience and there remained a continuing and substantial deficiency of establishment. The shortage of technical staff was exacerbated by a policy of leaving posts open, i.e. unfilled, for staff who left (the client's organisation) to work overseas.

Consultant organizations

Counterpart problems are by no means restricted to problems in the client organization, as the following report extract shows:

A central difficulty, never resolved, was the staffing of technical assistance consultants. The initiation of the project was delayed for over a year. The consultant staff originally proposed for the project made other commitments. Low consultant salaries plus the worsening economic conditions (locally) made it difficult to attract qualified experts. Since the elements of the programme were inter-related, delays in filling positions and poor performance in those positions made it certain that the balance of the programme would be affected.

Sometimes, although consultant staff may be technically competent, they lack the personal skills required to transfer technology in a counterpart situation. However, in many cases, consultants are under substantial commercial pressures to complete the project, and counterpart training is viewed both by them and the client as superfluous to this end. Many project time scales are simply totally inadequate to meet physical project objectives and to train staff.

Adapted from: Robinson (1988).

External training

External training may be provided in a national training institute, in a regional training centre operated by several countries, or in an overseas training centre. The advantages of external training are that more training resources can be made available, and that exchange of ideas from different contexts can be made when people from other domestic or foreign road organizations participate in the training.

A national training institute may either be an engineering college, a management institute, or a training centre geared specifically to the road sector's training needs. An example of a regional training centre is the Eastern and Southern Africa Management Institute (ESAMI) in Tanzania, which is operated by Tanzania, Kenya, Uganda, Zimbabwe, Djibouti and Comoros. ESAMI offers courses in international contracts and planning of road maintenance. Participants come from member and other countries.

National/
regional
training

Donors sometimes offer out-of-country training. This could either be post-graduate courses of 1 or 2 years duration at universities, short courses offered by the national roads administrations, or short fellowships tailor-made to meet specific skills requirements. Out-of-country training is sometimes preferred by the bilateral donors to promote cultural exchange between the countries. This type of training is also attractive to many trainees as it broadens their experience and develops maturity. The training is aimed mostly at medium to higher skill levels, including engineers, technicians, surveyors, accountants, economists, etc. This type of training is relatively more expensive than local training, but may be the only solution if the necessary training staff and training facilities cannot be provided nationally.

Out-of-
country
training

Training has been shown to be necessary as an integrated part of most development projects, if the projects are to succeed. For example, provision of new workshop facilities for repair of road equipment must include training of the mechanics who are to use the new facilities; introduction of a new road management system must be accompanied by systematic training of all the personnel who will be operating the new system. Training is particularly important for projects involving institutional strengthening or restructuring. However, experience from many projects suggests that training needs are often underestimated by a considerable amount (Robinson 1988). Project-related training tends to be very specific in its objectives as it is usually being delivered to attain well-defined performance requirements. It is likely to be of a short-term nature.

Project-
related
training

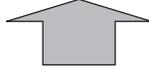
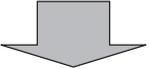
The twinning of organizations with similar but more mature organizations in other parts of the world has proved to be an effective way to transfer knowledge, train staff and build management capabilities (Cooper 1984). Professional relationships between operating entities offer advantages of being complementary and are flexible over time. The entity supplying technical assistance uses its own resources to offer services to its twin in the functional areas in which they both operate. Twinning arrangements in the roads field have been carried out, for example, between the US Federal Highway Administration and the Russian road administration. It should be noted that entities are not necessarily compatible simply because they are in the same business. Where twinning arrangements have been successful, there has always been an initial commitment and consensus on goals of the technical assistance. Both client and supplier may need help in setting up a twinning agreement and working together to carry it out. A tripartite arrangement between a host organization and an offshore institution in conjunction with a consultant seems to offer the best combination of skills and resources.

Twinning

The relevance of internal and external training and the appropriate type of training will vary according to training objectives and target groups. Often a blend of different type of training will be most relevant. There will be a wide variety of issues and methods needed within the operations of any road organization. The development of expertise and skills are likely to require a flexible approach to training, given the different backgrounds of staff involved. The concept of *training levels*, illustrated in Box 23.2, is useful for planning effective training that recognizes the requirement for

Training levels

Box 23.2 Training levels

Level 1: <i>Appreciation</i>	Broad awareness needed of the range of issues, methods and skills needed in the particular road administration.	Less understanding	More Staff
			
Level 2: <i>Knowledge</i>	Understanding of some detail of the principles involved in the methods and skills, and how they are applied in practice.		
Level 3: <i>Experience</i>	'Hands-on' experience in the use of the methods and skills for a particular function.		
			
Level 4: <i>Ability</i>	Detailed understanding and experience of the methods and skills needed in particular areas.	More understanding	Fewer Staff
Adapted from: Institution of Civil Engineers (1992).			

flexibility. As the training level increases, there is a corresponding increase required in the depth of knowledge of issues and methods, but a decrease in the number of staff who need to be trained. Training courses can be designed at a particular level that assume competence at the preceding levels. In this way, the training programme can be well focused and avoid repetition. At the first level of 'appreciation', training is likely to be more of a dissemination exercise.

Continuing
training

Training is not a once-only activity. There will be a need for on-going training as staff are replaced, as new methods are introduced and as skill requirements change. There is also a need for continual reinforcement of good practice through refresher training. It is often appropriate to train staff who are capable of undertaking future training activities themselves, without outside assistance. These staff will need not only to be motivated, but also to have appropriate interpersonal skills.

23.4 Training needs analysis

23.4.1 Determining needs

Identification

The demand for training needs to be identified by the road organization. As noted earlier, training needs analysis, carried out as part of an institutional appraisal, is the most effective method of identifying training objectives and institutional constraints. The analysis should comprise a determination of

- working environment, as affected by the institutional context;
- current performance levels;

- required performance levels; and
- training needs (the difference between current and required performance levels, as influenced by the working environment).

The current performance level has to be established for each category of target groups or position in the road organization, and this must be evaluated in relation to corporate objectives and performance requirements. The current performance level is evaluated by interviewing representative sections of the personnel and by reviewing job descriptions. Figure 23.1 is an example of a staffing and skills inventory form. Figure 23.2 shows a job description for a district maintenance engineer in Nigeria. It is also important to evaluate the performance level that can be expected from future employees. This may demand an analysis of existing schools and training institutes. Figure 23.3 is an example of a questionnaire that has been used for evaluating technical schools in Zimbabwe.

Current performance

The detailed training needs for each category of personnel are then assessed by comparing current and required performance levels.

Training needs

23.4.2 Training priorities

The training needs analysis is likely to identify more training than can be afforded, and decisions must be made about which to undertake now and which to defer. The aim should be to set training priorities on the basis of cost–benefit principles: the

Cost–benefit approach

Staffing and skills inventory form

Name of employee:
 Personnel no:
 Age:
 Job title:
 Salary grade:
 Length of service:
 Years in present post:

Normal duty:
 Exceptional duty:
 Work supervised by:
 Staff under supervision:

Promotion potential: • Supervisor's opinion
 • Management's opinion

Education level:
 Certificate obtained:
 Post-school education:
 Certificate obtained:

English: spoken : good/adequate/none
 reading : good/adequate/none
 written : good/adequate/none

Skills: • in which proficient
 • others

Figure 23.1 Staffing and skills inventory form.

Source: Cowi.

TECHNICAL OFFICER	(Periodic maintenance)
Place of work	Office and field
Work supervised by	District Maintenance Engineer
Staff under supervision	Technical Assistant and Specialist Overseers
Normal duty	<ul style="list-style-type: none"> • Instruction of Road Overseers • Supervision of work and reporting • Planning works programmes • Control of production units • Ordering materials • Allocating staff • Allocating machines and vehicles
Exceptional duty	Emergency works
Details	<ul style="list-style-type: none"> • Resurfacing • Surface dressing • Gravel production • Operation of quarry • Bridge and culvert repairs, etc. • Maintenance road furniture • Minor works

Figure 23.2 Job description for a district maintenance engineer in Nigeria.

Source: Kampsax.

choice of training to be undertaken should be that which gives the greatest return on investment. All potential training should be evaluated in this way. Examples of the use of a cost–benefit approach to prioritizing training have been given by Jorgensen (1988). These have been used as the basis of the example in Box 23.3.

23.4.3 Complementary requirements

Not all technical capacity problems can be solved by training alone. When corporate performance standards are established, complementary interventions to training are likely to be required to reach desired standards. As stated earlier, there is a high degree of correlation between training issues and institutional issues. But it may also be necessary to introduce new technologies and new management systems. The training should not be seen in isolation. If complementary requirements are not met, the intended change may not occur.

23.5 Planning

23.5.1 Basis for preparation

A plan for implementation of training should be prepared based on the definition of objectives, the training needs assessment and the priority analysis. This plan

1	SCHOOL Name: Location:
2	EDUCATION Content of curricula and syllabuses: Quality: Teaching methods: Examinations:
3	STUDENTS Entry qualifications: Guidance of students: Staff/student relationship:
4	STAFF Staffing policies: Selection and qualifications: Staff development and training:
5	PHYSICAL RESOURCES Teaching accommodation: Provision of equipment:
6	RELATIONS with construction and maintenance units in the field
7	ORGANIZATION Organization and structure: Objectives and plans:
8	EFFICIENCY Admission rate/drop-out rate: Time to produce a graduate: Staff/student rate: Accommodation utilization: Costs:
9	NATIONAL SYSTEM Central policies: Control of courses: Control of staff: Sources of funds:

Figure 23.3 School questionnaire.

Source: Cowi.

should specify:

- Criteria for selection of trainees
- Type of training to be used
- Training programme.

A number of factors influence the impact that can be achieved by training interventions. These factors may be termed the training environment and are concerned with supply, mobility and trainability of personnel. It is important to have an understanding of the training environment as it will have to be incorporated in the design of the

Box 23.3 Example of prioritizing training needs

Consider the following situation concerned with the grading of a network of unpaved roads: (i) each grading of the road is carried out using six passes of a grader; (ii) but the graders are only available for use in operational condition for 20 per cent of the normal working hours; (iii) as a result, 75 graders are needed for maintaining the network of roads. A training programme for grader operators, designed to teach them an improved work method, which can produce the same result using only four passes of the grader, results in a saving US\$1.25 million. Alternatively, if improved equipment maintenance can increase the availability rate to 50 per cent, the number of graders needed can be reduced from 75 to 30, with an estimated saving of US\$2.25 million. If both operators and mechanics are trained, the total saving will be US\$2.75 million. The results are summarized below. Although these examples are dramatic in their impact, this approach to evaluating the benefits of training can and should be applied to all activities in order both to justify and to prioritize training needs.

Alternative grader fleet sizes under four different scenarios of work method and availability rate

	<i>Scenarios</i>			
Work method	6 passes	4 passes	6 passes	4 passes
Availability rates	20%	20%	50%	50%
Required fleet size	75	50	30	20
Fleet investment ^a (\$US millions)	3.75	2.50	1.50	1.00

Note

a Assumes US\$50,000 per grader.

training intervention if the desired impact is to be achieved. Any planning must recognize the environment in which the training will be carried out.

23.5.2 Training environment

Supply of staff

The supply and mobility of staff are strongly influenced by the prevailing education system in the country. An elitist education system, which concentrates resources on higher education for the few, tends to have a lack of middle-level technicians and skilled workers – in contrast to an education system, which aims at a more appropriate balance of primary, secondary and higher education. The geographical dispersion also affects the availability of staff.

Mobility

Higher educated personnel often have a preference to remain close to the educational centres where they received their education. As these institutions are normally concentrated in the capital, or in a few large cities, it may prove difficult to attract higher skilled personnel to more remote areas. Incentives can sometimes be provided to increase mobility. There will normally be little scope for substantial salary increases, so other

types of incentives have to be considered. These might include improved career prospects, a challenging work environment, housing, etc. Provision of these may increase the mobility to the more remote areas.

Trainability is concerned with the extent to which training can make an impact on staff's attitudes, qualifications and skills. It relates to quality of education and work experience of the individual staff member. It may also relate to age, since younger staff tend to be more flexible. However, this does not mean that more mature staff should be excluded from training opportunities. The type of training will have to be adapted to the individual's absorptive capacity for theoretical and practical training.

The above factors will generally be outside the control of the road organization. But there may be some scope for influencing them through project activities. For example, a project-related training intervention may provide an attachment to a training or education institution dealing with road engineering. This could assist in modifying educational curricula.

Trainability

Influence

23.5.3 The trainees

The selection of trainees will normally be undertaken by the road organization according to established criteria. However, in practice, the criteria are often not very relevant and are not always adhered to strictly. Modifications may be necessary to the normal approach to make training appropriate to all participants. When selecting trainees for out-of-country training, special measures may be taken to assure that the necessary qualifications have been obtained. A representative from the donor may interview candidates before final selection. Language ability will often be a requirement.

Engineers and other professionals are used to classroom situations, but training should never focus on lectures only (one-way communication). The trainees should be enlivened through discussions, and exchange of experiences should be encouraged. Classroom training should also be combined with field visits in order to illustrate the taught theoretical aspects.

Skilled and unskilled workers are normally not used to acquiring knowledge through classroom training or by reading books. Nor are they used to being in a learning situation for long periods. Training of manual workers should, therefore, focus on practical on-the-job training. Classroom training should take up no more than 40–50 per cent of the time. Audio–visual aids such as film, slides, videos, overheads and models should be used extensively, although this applies to all training. A training course for manual workers should have a duration of no more than five weeks. The course should be divided into modules if training of longer duration is necessary. The trainees should be sent back home between the modules to practise their newly acquired skills. This approach is called 'sandwich training'.

Selection of trainees

Engineers

Manual workers

23.5.4 Training programme

The training programme (curriculum) should be based on the identified needs of target groups. The programme should give a statement of which courses should be arranged for the different categories of personnel. Further, the programme should state the objective, the content and the duration of the individual courses. Figure 23.4 shows a training programme prepared for the road organization in Sabah, Malaysia

Targeting need

Sabah Public Works Department
Development and staff training programme

Trainees		Course description	1984											
Group	Number		Jan.	Feb.	March	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Engineer Trainee	1	Instructor Course A	20											
Instructors Trainees	4	Instructor Course B	20				4	10						
Overseers & O's trainees	12	Road Maintenance I	6		17	2	11							
Overseers & O's trainees	12	Road Construction I					4	30						
Technical Ass. & T.A.'s trainees	12	Road Maintenance II							13	22				
Technical Ass. & T.A.'s trainees	12	Road Construction II									15	8		
Overseers & Technical Ass.	8	Surface Dressing					14		13					
Overseers & Technical Ass.	8	Asphalt Concrete								24	13			
Overseers & Technical Ass.	12	Aggregate Production						2	21					
Engineers & Eng. Ass.	20	Cost control & Admin.			15	23								
Engineers & Eng. Ass.	25	Review					15	18						11 25

Figure 23.4 Training programme from Sabah, Malaysia.

Source: Kampsax.

<p>OBJECTIVES To train the officers in basic road planning aspects.</p> <p>ORGANIZATION The course programme should be elaborated by the same working groups as the course for PPOs. The working groups should be assisted by one or two Provincial Road Engineers.</p> <p>DURATION: 15 days</p> <p>NUMBER OF COURSES: 10</p> <p>TOTAL NUMBER OF PARTICIPANTS: 110</p> <p>LOCATION The most conveniently situated Provincial Training Centre run by the Public Services Commission or the nearest DDF training centre.</p> <p>IMPLEMENTATION The coming years in step with the amalgamation of Rural and District Councils.</p> <p>COURSE CONTENT</p> <ul style="list-style-type: none"> • Basic mathematical concepts • Elementary map reading • Road standards • Basic feeder road construction and maintenance • Assessment of the technical viability of proposed new roads • Calculation of cost estimates for construction projects • Identification of road maintenance needs • Benefit of feeder roads, e.g. induced agricultural production and reduced transport costs.

Figure 23.5 Description of a course for senior field officers in Zimbabwe.

Source: Cowi.

and Figure 23.5 is a syllabus for a course planned for senior field officers in Zimbabwe.

The need for training will vary considerably depending on the target group's skill and qualification levels. For example, the need for skills and qualifications will be different in a national and centralized trunk road organization, than in a local and decentralized rural road organization. In general, planning involves more than identifying a topic, a time and a date. The content and form of a course can be considered in terms of a number of issues, some of which are listed in Box 23.4.

The training programme should be presented to the executive management of the road organization in order to get their approval and positive support. The presentation can often take the form of a 1 or 2-day seminar to allow for a thorough discussion and possible amendment. The presentation may also assist in identifying areas for management training to be incorporated in the overall training programme.

Target group

Presentation

23.5.5 Training consultancies and contracts

Often, project-related training is carried out by consultants under contract. The training programme to be planned may relate to a number of different road activities. Sometimes the various training activities are contracted out to different firms, but

Firms undertaking the training

Box 23.4 Determining the form and content of a course*Form of course*

- What length will the course be?
- Should it be full-time or part-time?
- Should it be residential or non-residential?
- How many sessions per day?
- What kind of sessions each day?
- What should be the balance between lecturing, instructing, demonstrations and discussions?
- What is the maximum length of unbroken lectures?
- Should the previous day's work be reviewed each morning?
- What use should be made of audio/visual aids, models and other training material?

Course content

- What are the precise objectives?
- What is the overall strategy?
- Will the final concepts be defined at the outset?
- Who will write the introductory outline?
- How will progressive learning be achieved?
- What will be the pattern and the structure?
- What methods and techniques will be used?
- How will maximum course member participation be achieved?
- How and when will the course tutors be briefed?
- Should a synopsis for each session be given or agreed?
- What further help will tutors need?
- How can scope be left for amendments?
- How is it possible to ensure flexibility, continuity and integration?
- Should a course tutor always be present?
- How best can the tutor act as a catalyst?
- What provision should be made for free discussion?
- Is it necessary to prepare handouts?
- Is it necessary to have worked examples?
- When and what periods should there be for private study?
- How can the optimum balance of activities be ensured?
- Is it necessary for tutors to have 'dummy runs' or rehearsals; if so, when?
- What is the purpose and what should be the frequency of outside visits?
- What period should be allowed for final preparations?
- What new sessions should be included?
- What revised timings should be made?
- What open forums are necessary for course members?

Adapted from: Barber (1968).

this is likely to cause conflicts and complications. The training may be awarded to a training consulting firm, which would then be responsible for all of the training inputs, and may sub-contract appropriate courses to be undertaken by consulting firms and contractors. In these cases, the following approach to planning is usually followed.

A broad outline and concept of the conceived training intervention is described in the terms of reference (ToR). These are normally prepared by the road organization, by a donor or may be the result of a study undertaken by a consultant. The ToR should state the training objectives, expected outcomes, as well as course content. Advice on drafting ToRs is given by Lewis (1998).

Terms of reference

The ToR forms the basis for preparing the technical and financial proposals by the firms bidding for the assignment. Since the proposal establishes the framework for conduct of the training, it is important already at this stage to be aware of the training environment, in addition to the specific training requirements. The training environment may not be described adequately in the ToR so, to reduce uncertainty, the firm submitting the proposal often makes an initial survey to ensure that what is offered in the proposal is compatible with the prevailing situation in the project area. Advice on preparing proposals is given by Lewis (1992).

Proposals

23.6 Detailed preparation

The detailed preparation of the training programme can start immediately after the plan of implementation has been approved. This phase of activity includes

Activities

- selection of trainers;
- preparation of lesson plans;
- procurement and production of training aids;
- provision of training facilities; and
- planning any evaluation of the training that will be required.

Trainers will often come from a road organization's own training school, or from a local training institution, in which case they will usually have the required skills and be familiar with the training environment. However, in the case of project-related training, the project team members are frequently professionals and technicians without specific teaching and training expertise. They will have been chosen principally because of their know-how and experience. It may therefore be important to provide them with some training themselves in order to facilitate the detailed preparation of the lessons; how to communicate with the trainees; how to prepare teaching aids, etc. In the event that the training inputs are substantial, it is advisable to arrange a course for the team members in teaching techniques. In most cases, it is usually best to select individuals who have a balance between technical knowledge and training skills.

Selection of trainers

Professionals and technicians who are going to train people from other countries need to meet some particular requirements. First, they should be proficient in the language of instruction. This requirement seems self-evident, but is not always met. Second, the trainers ought to have some experience from practical work in the home country of the trainees. Otherwise it is difficult for the trainers to relate the training

Foreign trainers

to the needs of the trainees. Added to this is the need to recognize that association with people from foreign cultures can involve sociological problems that may be difficult to cope with for inexperienced trainers. This can present particular problems because great authority is normally attached to a trainer.

Lesson plans

Detailed lesson plans will have to be prepared for each course based on the curriculum. A lesson plan consists of a timetable describing the content of each lesson and type of training (classroom lecture, group work, practical training, etc.). The lesson plans are often prepared by, or in co-operation with, the individual trainers who are going to deliver the training. It is important that the lesson plans are prepared carefully to prevent overlap and to support inexperienced trainers. Figure 23.6

SEMINAR TIME SCHEDULE 研讨课程的时间表	
简例： Legend: TW = Tim Waage 添·华治 MA = Magne Aasen 马·阿森 BT = Bent Thagesen 宾·泰捷逊 DRD= Danish Road Directorate 丹麦公路局	
Monday 30/6 1986	1986年6月30日星期一
8.30 - 9.00	Opening Ceremony 开课仪式
9.00 - 9.30	Introduction to Seminar (BT) 介绍研讨课程 (BT)
9.30 - 10.30	Introduction to Kampsax, Slide show (TW) 介绍金硕, 放映幻灯片 (TW)
10.30 - 11.00	Video: "On the Way" (TW) "在路上" 录像带 (TW)
11.00 - 12.00	Introduction to DRD (MA) 介绍丹麦公路局 (MA)
14.30 - 17.30	Lecture A.1 - "How to Execute a Project" (TW) 讲章 A.1 - "如何执行一项工程" (TW)
Tuesday 1/7 1986	1986年7月1日星期二
8.00 - 9.30	Group work for Lecture A.1 讲章 A.1 的小组工作
9.30 - 12.00	Lecture A.2 - "Introduction to FIDIC Contract Documents" (TW) 讲章 A.2 - "介绍国际顾问工程师联会的合同文件" (TW)
14.30 - 15.30	Lecture A.2 (continued) 讲章 A.2 (续)
15.30 - 17.30	Group work for Lecture A.2 讲章 A.2 的小组工作
Wednesday 2/7 1986	1986年7月2日星期三
8.00 - 12.00	Lecture A.3 - "Bill of Quantities and Schedules" (MA) 讲章 A.3 - "工程量清单和明细表" (MA)
14.30 - 16.30	Group work for Lecture A.3 讲章 A.3 的小组工作
16.30 - 17.30	Lecture A.4 - "Selection of Engineer and Contractor" (TW) 讲章 A.4 - "选择工程师和承包商" (TW)

Figure 23.6 Lesson plan for a training course for road engineers in the Peoples Republic of China.

Source: Kampsax.

shows a lesson plan for a training course for road engineers in the Peoples Republic of China.

Training aids consist of course notes and audio–visual aids. Many commercial training aids are available, but they are expensive and often not very suitable for specific purposes because they are too general. Usually, it is better to produce course notes, slides, videos, etc. that are specific to the detailed training being undertaken. This makes it possible to illustrate local issues and problems. It is important to reserve sufficient money for procurement and production of training aids when preparing training proposals.

Training aids

Course notes should be in the language of the trainees. It is common for these to be of two types:

Course notes

- paper copies of visual aids, which can be used by participants during the training, and can be annotated with notes and comments; and
- detailed written notes covering the technical content of the training, which can be studied after the training has been completed, and used as a source of reference.

The presentation in the course notes should be tailored to the trainees' capabilities. Course notes for professionals may contain a great deal of plain text. Course notes for slow readers may be better presented as strip cartoons, as in the example in Figure 23.7.

The preparation of visual aids requires some care. Good slides amplify and clarify the message, stimulate interest and help the speaker keep 'on track'. They merit the same care in preparation as the text of a presentation. The primary consideration is *legibility*, since slides that cannot be read when projected lessen the impact and effectiveness of the presentation. Copies of typescript, computer outputs, graphs and illustrations from textbooks are normally too small to be seen when projected on to a screen. As a rough guide:

Visual aids

- no slide should contain more than 15–20 words and
- the type size on overhead projector slides should be a *minimum* of 7–10 mm in height; a minimum font size of 18 pt should be used in Powerpoint presentations.

Graphs and drawings usually need to be redrawn with increased line widths, and with captions reduced in number, simplified and increased in size. Each slide should cover only one idea: it is better to use several simple slides rather than one complicated slide.

Depending on the type of training, the required training facilities may include classrooms, workshops, machines, accommodation for the trainees, etc. In some cases, the training facilities already exist and are readily available; in other cases, the facilities have to be constructed and furnished as part of the training project.

Training facilities

Box 23.5 provides a checklist of some of the items that need to be considered during the preparation for training courses.

Course administration
Evaluation and tests

Evaluation procedures and tests should be developed in advance, during the preparation stage, to assess the extent to which the training objectives have been met. Evaluations are normally applied to judge qualifications and attitudes. A training evaluation pack needs to be developed for each session. This will consist of both test material and evaluation criteria. Tests are applied to judge specific skills, and simple



Figure 23.7 Course note from a training course for road overseers in Sabah, Malaysia.

Source: Kampsax.

test material should be provided that could be used during and at the end of a session to test comprehension. This can be written, but is normally oral, for use in an interactive manner. Further evaluation material is normally provided for use at regular intervals after the formal training has been completed to ensure that principles and methods

Box 23.5 Checklist for course preparation

- Ensuring initial discussion of details with all concerned
- Drafting the initial programme
- Securing line management approval of the draft
- Producing a final well laid out programme
- Passing final confirmation to all speakers
- Securing agreed lists of course members
- Notify course members well in advance
- Ensuring adequate accommodation
- Arranging adequate seating, lighting and ventilation
- Giving clear instructions on contact arrangements
- Ensuring appropriate breaks and refreshments
- Having adequate power points and black-out, if needed
- Securing required equipment, models and materials on time
- Arranging for any required bench space or outdoor work areas
- Arranging for required display facilities
- Making contingency arrangements for 'stand-in' sessions
- Organizing the payment of fees and expenses
- Making arrangements for meeting and introducing visiting speakers
- Coping with last minute changes to the timetable
- Checking and distributing handouts
- Arranging typing and copying
- Arrangements for any 'feedback' to speakers
- Arrangements for any course reporting
- Arrangements for discussions of course members with supervisors.

Adapted from: Barber (1968).

are still understood, and to identify any needs for retraining. Evaluation procedures may also be developed to obtain the trainees' own impression of the quality of the training provided. Diplomas and certificates can be provided to successful candidates.

23.7 Implementation

23.7.1 Communication

Communication is the essence of effective training. Without effective communication, no instructions can be given or received and nothing can be achieved. Communication is discussed in Box 23.6. There are several modes of communication, each with its own set of mechanisms and channels:

- Oral
- Visual
- Written.

Modes of
communication

Box 23.6 Communication

To have effective communication, it is necessary to understand the basic mechanisms. Communication is more than just ‘telling’; it has three components:

- Source
- Channel
- Receiver.

In order for information to pass from the source (trainer) to the receiver (trainee), the following things must take place:

- first, the trainer must know, with some precision, what is the message to be conveyed;
- second, the message, which consists of ‘internal’ thoughts, must be ‘externalized’ through the use of both verbal language and ‘body’ language;
- third, the trainee must have excellent reception of the signals being transmitted; and
- fourth, the trainee must integrate the numerous signals received and turn them into understandable thought and action.

The quality of reception and understanding will be conditioned by the trainee’s

- prior knowledge and basic assumptions;
- attitude;
- frame of mind.

A breakdown or inadequacy in any of these steps will prevent effective communication taking place.

In the context of roads training, the oral and visual modes are those most used, although written communications also have a role. Effective communication is more difficult to achieve than most people realize, and success rates are typically less than 25 per cent (Decker 1988). This means that the odds are 3 : 1 against communication being successful, with the message received being the one that the trainer intended to convey.

Feedback

It is important that the trainer is aware of these constraints, and addresses them when framing the way that the information is transmitted. In order to check that communication has taken place, and that information has been received, understood and produced the required reaction, some form of feedback is advisable. Without effective feedback, the effectiveness of future training may be hampered through a lack of appreciation by the trainer of the reasons for ineffective communication. It is, therefore, crucial that, for training to be effective, the trainer understands the basic mechanisms of communication.

23.7.2 *Instruction*

It is beyond the scope of this book to provide a detailed training manual. However, Box 23.7 provides some examples of where inexperienced trainers offend the most basic rules of good instruction.

Examples

23.7.3 *Monitoring and feedback*

To ensure that training objectives are being met, on-going monitoring will be required, and this will be used to refine the training programme. Personnel responsible for training can use tests and questionnaires for participants for this purpose, in addition to any assessments. Monitoring should take place at regular intervals through the training.

Monitoring

The evaluation made at the end of the training programme should reveal the extent to which objectives have been met. Where objectives have not been met, feedback mechanisms need to ensure that subsequent training activities are modified to avoid repeated mistakes. However, the progress monitoring should prevent any serious problems arising.

Feedback

Even in those cases where the training objectives have been met, this does not necessarily mean that the corporate performance requirements will also be achieved. For example: the road organization may not draw on the newly acquired qualifications of the trainees; some trainees may not want to share their new knowledge with others. Issues such as these can only be revealed after some time when the trainees

Post
evaluation

Box 23.7 Common problems when instructing

- The trainer is not properly prepared; the fact that a good instructor never reads direct from a manuscript sometimes leads inexperienced instructors to believe that a lecture does not need to be prepared carefully; however, even experienced instructors attach much importance to the preparation of lectures.
- The trainer speaks to the white board or to the screen instead of facing the audience; the trainer should always be aware of the audience.
- The size of the lettering on slides is so small that it cannot be read from the back of the room.
- The trainer gives too much and too detailed verbal information; lectures are only suitable for presenting a general view of a subject and for emphasizing a few essential messages.
- The trainer focuses too much on one-way communication; discussions, exercises and group work, should play an important part in training; however, preparation of this type of training requires experience and is time consuming.
- The duration of classes is too long; it has been demonstrated that the benefit from a lecture decreases rapidly with the duration of the lecture: a lecture should never last more than 30–40 minutes without a break.

have been working in their assigned positions. A post-evaluation is seldom made, but it can provide a rich source of valuable information that could be utilized either in the organization's permanent training set-up, or in the design of succeeding training programmes.

References

- Barber, J.W. (ed.) (1968) *Industrial Training Handbook*, London: Iliffe.
- Brooks, D.M., Robinson, R. and O'Sullivan, K.P. (1989) 'Priorities in improving road maintenance overseas: a check-list for project assessment', *Proceedings of the Institution of Civil Engineers, Part 1*, **86** (Dec), 1129–1141.
- Cooper, L. (1984) 'The twinning of institutions: its use as a technical assistance delivery system', *Technical Paper No. 23*, Washington, DC: The World Bank.
- Decker, B. (1988) *How to Communicate Effectively*, London: Kogan Page.
- Institution of Civil Engineers (1992) *Management Development in the Construction Industry: Guidelines for the Professional Engineer*, London: Thomas Telford.
- Israel, A. (1987) *Institutional Development: Incentives to Performance*, Baltimore: Johns Hopkins for the World Bank.
- Jorgensen, J.S. (1988) 'Training in the transportation sector: a necessity', *National Development*, March, 36–41.
- Lewis, H. (1992) *The Consultant's Complete Proposal Manual*, London: PTRC Research and Education Services.
- Lewis, H. (1998) *Terms of Reference*, London: PTRC Research and Education Services.
- Moeller, P.W. and Iacono, M. (1990) 'Human resource and institutional development in the road sector', *Sub-Saharan Africa Transport Program. Human Resource and Institutional Development Project: Modal Summary Report*, Washington, DC: The World Bank.
- Nadler, L. and Nadler, Z. (1989) *Developing Human Resources*, 3rd edn, San Francisco: Jossey-Bass.
- Relf, C. and Thruscutt, S. (1991) 'Human resource development and management', in EDI *Sub-Saharan Africa Transport Program, The Road Maintenance Initiative, Volume 2: Readings and Case Studies*, Washington, DC: The World Bank, 178–181.
- Robinson, R. (1988) 'A view of road maintenance economics, policy and management in developing countries', *Research Report 145*, Crowthorne: Transport and Road Research Laboratory.
- Robinson, R. (1991) 'Institutional development in road maintenance organizations', in EDI *Sub-Saharan Africa Transport Program. The Road Maintenance Initiative, Volume 2: Readings and Case Studies*, Washington, DC: The World Bank, 168–177.

Development aid

Jens Erik Bendix Rasmussen and Tonny Baek

24.1 Introduction

The concept of development aid has changed over time, as described in Chapter 1. The United Nations (UN) and the World Bank have played a leading role in developing aid concepts and in streamlining procedures for preparation and monitoring development aid activities. In recent years, the World Bank and the International Monetary Fund (IMF) have expanded their economic reform programmes to include debt relief and to request recipient countries to formulate poverty reduction strategy papers (PRSPs). These provide a common operational framework for addressing poverty reduction and for achieving the internationally accepted ‘Millennium Development Goals’ (see Chapter 1) by 2015.

Aid concept

Thus, a key reason for rich countries to provide aid is to contribute to poverty reduction in poor countries and reduce inequality. It is widely recognized, however, that the achievement of this depends on a number of factors, including institutional improvements, less corruption, peace and stability, improved market access in the rich countries – especially for agricultural products. In addition to the poverty reduction motive, emergency assistance for purely humanitarian reasons plays a role.

Reasons for providing aid

There is a growing general acceptance among donors and low-income countries that deficiencies must be tackled simultaneously in a co-ordinated and comprehensive framework and partnership using country PRSPs as a guiding tool. Even if the Millennium Development Goals are not binding for the donors, it is expected that their existence will inspire some donors to raise their levels of aid in the coming years.

Approach

24.2 Resource transfers

Official development assistance (ODA), or aid, is defined as a resource flow that includes a grant element of at least 25 per cent. This element may be a ‘gift’ that does not have to be repaid, or the resource flow may be a loan with an interest rate or repayment period substantially more favourable than could be obtained on normal commercial terms. ODA is therefore concessional in character. ‘Official development finance’ (ODF) includes all ‘official’ resource flows, including those where the concessional element is less than 25 per cent.

Aid

Over the years, the willingness of rich countries and their taxpayers to assist poor countries has varied and different motives have been at play. There have been ideological motives, especially during the cold war, commercial interests, historical

Levels of aid

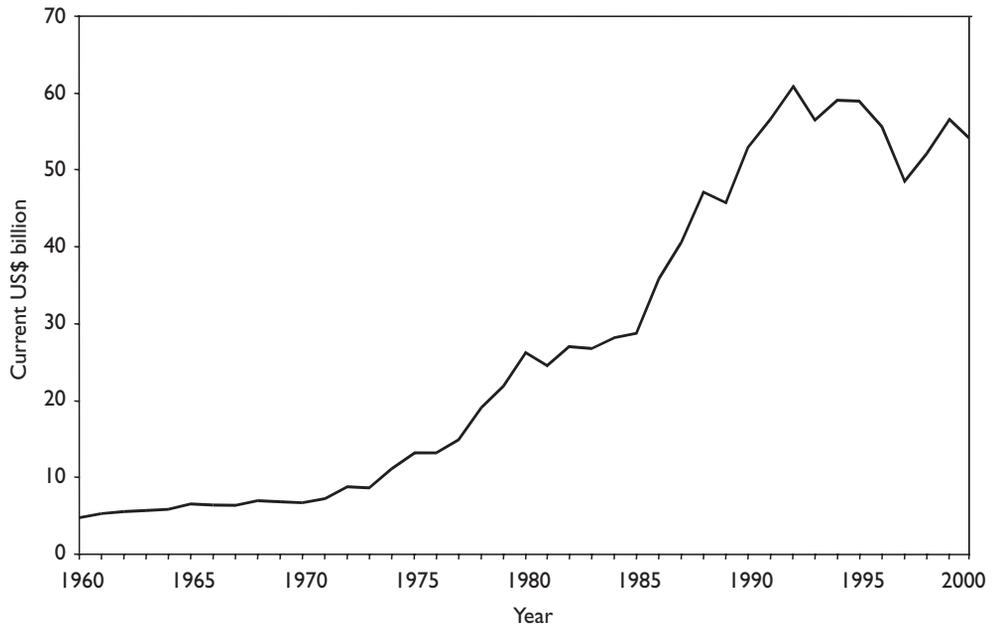


Figure 24.1 Net ODA transfers.

Source: OECD (2002).

colonial ties and, more recently, the potential effects on peace, stability and democratic rule and the impact of terrorist activities. The level of aid has been affected negatively by the difficulties of achieving sustainable growth and development in a number of countries, especially in Africa. The lack of progress or, in some cases, even decline have been due to different factors, such as weak institutions, bad policies, mismanagement, corruption, instability and war. In other cases, environmentally unsustainable activities have undermined long-term development. The fluctuations in ODA over time are shown in Figure 24.1.

Importance
for developing
countries

Donor support to national development efforts in different sectors, including the road sector, plays an important role in many developing countries. The amount of donor support varies greatly between countries, as illustrated by Table 24.1, which shows ODA as percentages of gross national income (GNI) for some selected countries.

Multilateral/
bilateral aid

Development aid is normally divided into 'multilateral' and 'bilateral' assistance. Multilateral assistance is aid transferred from a donor country via international aid organizations. Bilateral assistance is aid transferred directly from the donor country to a recipient country.

Assistance
from DAC
countries

Prior to 1989, the Eastern Bloc provided a sizeable share of the transfers from industrialized to developing countries. However, following the break-up of the Soviet Union, transfers from countries belonging to the Organisation for Economic Co-operation and Development (OECD) dominate the picture. Thus, total net transfers from the members of OECD's Development Assistance Committee (DAC) in 1999 were US\$251 billion, comprising ODF, export credits, and private transfers including private investments, bank and bond loans and assistance from non-government organizations (NGOs). Figure 24.2 illustrates that the net flow of private

Table 24.1 ODA for selected countries (1999)

Country	Percentage of GNI
Bangladesh	3
Bhutan	16
Mozambique	22
Nicaragua	34
Tanzania	11
Zimbabwe	5

Source: OECD (2002).

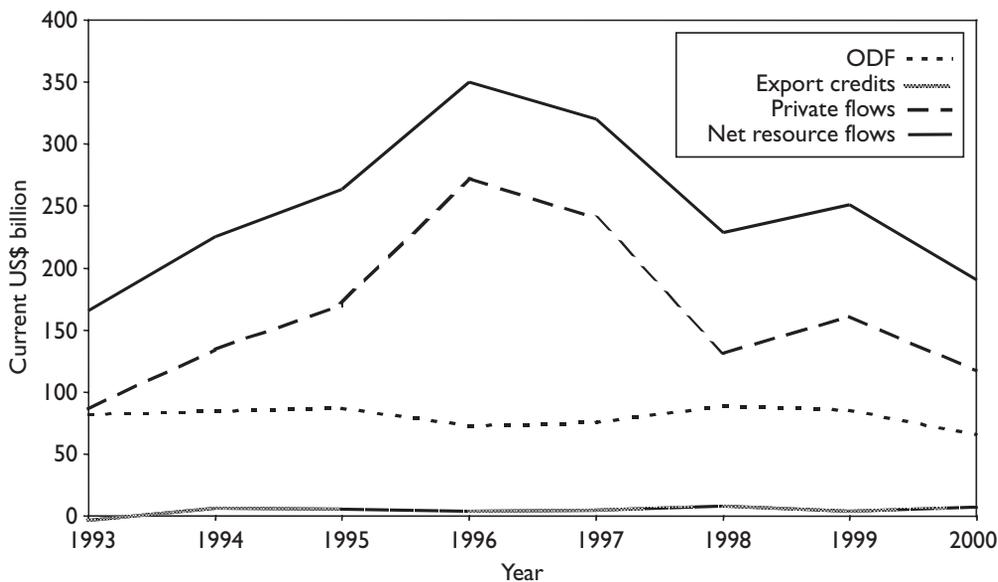


Figure 24.2 Net flow of ODA funds 1993–2000.

Source: OECD (2002).

funds have been significantly higher than the ODF during the period 1993–2000, but that resource flows have fluctuated widely with a peak in 1996. In Table 24.2, the transfers from DAC member states in 1999 are broken down in more detail. ODF amounted to a total of US\$86 billion, or 34 per cent of the total net flows. Bilateral transfers of US\$53 billion dominate, compared to US\$33 billion multilateral transfers. Private transfers amounted to US\$161 billion or 64 per cent of the total net flows.

Table 24.3 details the ODA from the DAC countries in 2000. In 1970, the UN set a target that international development aid should reach 0.7 per cent of gross national income (GNI). Over 30 years after this recommendation, the aid from the DAC countries amounts only to 0.22 per cent overall. Only five countries had reached and passed the target by 2000, with Denmark contributing the highest percentage. Japan makes the largest absolute transfers of funds. The ratio of bilateral aid to multilateral aid varies. On average, 33 per cent of the aid was multilateral.

Official
transfers from
individual
DAC
countries

Table 24.2 Net resource flows from DAC member states to the developing countries (1999)

Type of transfer	US\$ billion	%
ODF		
ODA	52.1	21
Bilateral	37.9	
Multilateral	14.2	
Official aid	7.8	3
Bilateral	4.9	
Multilateral	2.9	
Other	26.1	10
Bilateral	10.4	
Multilateral	15.6	
<i>Total export credits</i>	4.0	2
<i>Private flows</i>	161.0	64
Direct investment	145.6	
International bank lending	-79.6	
Total bond lending	28.8	
Other	59.5	
Grants by NGOs	6.7	
Total net resource flows	251.0	100

Adapted from: OECD (2002).

Table 24.3 ODA from the DAC countries (2000)

Country	US\$ million	GNI (%)	Bilateral (%)	Multilateral (%)
Denmark	1,664	1.06	62	38
Holland	3,135	0.84	72	28
Sweden	1,799	0.80	69	31
Norway	1,264	0.80	74	26
Luxembourg	127	0.71	74	26
Belgium	820	0.36	58	42
Switzerland	890	0.34	70	30
United Kingdom	4,501	0.32	60	40
France	4,105	0.32	69	31
Finland	371	0.31	59	41
Ireland	235	0.30	66	34
Japan	13,508	0.28	72	28
Germany	5,030	0.27	53	47
Australia	987	0.27	77	23
Portugal	271	0.26	65	35
Canada	1,744	0.25	67	33
New Zealand	113	0.25	75	25
Austria	423	0.23	61	39
Spain	1,195	0.22	60	40
Greece	226	0.20	44	56
Italy	1,376	0.13	27	73
United States	9,955	0.10	74	26
Total	53,739	0.22	67	33

Source: OECD (2002).

24.3 International aid agencies

24.3.1 Financial institutions

Multilateral aid is mainly transferred to developing countries through the World Bank Group, the regional development banks, the UN-system and the European Union (EU). Membership of international aid organizations comprises both developed and developing countries, with each country contributing to the operation of the organization. Members' contributions generally reflect their economic strength.

The most important multilateral financial institution is the World Bank Group. It includes the International Bank for Reconstruction and Development (IBRD), the International Development Association (IDA), the International Finance Corporation (IFC), the International Monetary Fund (IMF) and the Multilateral Investment Guarantee Agency (MIGA), all with head offices in Washington, DC. Practically all countries are members of the IBRD and IDA. The objective of IBRD is to provide financial assistance on terms more favourable than normal commercial rates. IDA provides long-term interest-free credits, which carry only a nominal service charge. IDA credits are only available to the poorer countries with a GNI per capita of less than US\$875 (2001). The IFC provides financial assistance for the development of the private sector. The IMF supports monetary stability and provides credits to improve balance of payments. MIGA has the objective to encourage foreign direct investment in developing countries by protecting investors from non-commercial risks, including the risk of war. The World Bank Group works worldwide.

The Inter-American Development Bank (IDB) has its head office in Washington, DC. Its objective is to support economic development in Latin America. The African Development Bank (AfDB) and its development fund (AfDF) have their head office in Abidjan, Ivory Coast, and operate in Africa. The Asian Development Bank (AsDB) and its development fund (AsDF) have their head office in Manila, the Philippines, and operate throughout Asia. Membership of the regional banks is normally by members of OECD plus countries in the region concerned.

The International Fund for Agricultural Development (IFAD), with head office in Rome, is a special UN fund for financing investment in the agricultural sector in developing countries.

The European Bank for Reconstruction and Development (EBRD) was established in 1991. The purpose of the EBRD is to help build market economies and democracies in 27 countries in Central and Eastern Europe and Central Asia.

The net amount of funds transferred (disbursements minus repayments) to developing and emerging countries from international financial institutions in 2000 is listed in Table 24.4. The financial institutions contribute to policy change through financial support to programmes and projects and, in the case of the World Bank, through structural adjustment lending.

Multilateral
aid

World Bank
Group

Regional banks

IFAD

EBRD

Transfers

24.3.2 The UN system

The UN forms an umbrella for more than 50 different specialized agencies. Some of these have activities in all member states including the developed countries. Others are purely development aid agencies, with activities only in developing and some

Agencies

Table 24.4 Net disbursements from financial institutions (2000)

<i>Institution</i>	<i>US\$ million</i>	
	<i>Concessional</i>	<i>Non-concessional</i>
AfDB		-304
AfDF	300	
AsDB		1,409
AsDF	927	
Caribbean Development Bank	20	50
EBRD	5	237
IBRD		2,776
IDA	4,188	
IDB	153	4,360
IFAD	143	5
IFC		229
IMF	-89	
Nordic Development Fund	38	
Total	5,716	8,402

Source: OECD (2002).

former Eastern Bloc countries. UN assistance is generally in the form of grants and technical assistance. Different specialized agencies implement both projects and programmes within their subject area.

Technical assistance and co-operation

Technical assistance (TA) and technical co-operation (TC) are typically used to assist a developing or emerging country in improving capacity of institutions over time. A typical example could be training of staff in a road administration in installation and operation of a 'pavement management system', which is a tool for planning and budgeting the optimal use of scarce funds for road maintenance (see Chapter 19). TA or TC may also be used to advise on sector policy improvements, regulatory frameworks, etc.

UNDP

The United Nations Development Programme (UNDP) is the central body for financing and co-ordinating UN's development aid. Twelve per cent of UN's development aid was channelled through UNDP in 2000. Projects financed through UNDP are normally implemented by specialized UN agencies.

UNICEF, UNHCR and ILO

Major UN implementing agencies are the United Nations Children's Fund (UNICEF) and the United Nations High Commission for Refugees (UNHCR). The objective of UNICEF is to improve the condition of life for children and mothers. The effort is concentrated on health, nutrition, water, sanitation and education. UNHCR is in charge of international aid to refugees. The aid from this agency includes emergency relief, development projects and education. The International Labour Organization (ILO) is particularly relevant to the roads sector because it is very active in promoting labour-based methods in road construction and maintenance.

Other UN agencies

The World Food Programme (WFP) provides food assistance to developing countries. More than half of the projects implemented are so-called 'food for work' projects, many of which include road construction and maintenance. WFP is also

Table 24.5 Net disbursements from UN agencies (2000)

Organization	US\$ million
UNICEF	576
UNHCR	493
UNTA	454
UNDP	390
WFP	357
UNRWA	301
UNFPA	133
Other UN agencies	568
Total	3,272

Source: OECD (2002).

distributing food as emergency relief to refugees. The United Nations Transition Assistance group (UNTA) has assisted Namibia and Cambodia in the transition to independence and democratic elections. East Timor and Kosovo have also been assisted. The United Nations Relief and Works Agency (UNRWA) is a special aid organization for the Palestinian refugees. The United Nations Fund for Population Activities (UNFPA) promotes family planning. The organization is mainly financing projects implemented by other UN organizations or by the recipient countries. Among other important UN bodies is the World Health Organization (WHO), which assists health programmes in the developing countries. The programmes include vaccination campaigns, setting up of systems for storage and distribution of medicine and AIDS control.

The total development aid from the UN system was US\$3.3 billion in 2000. The aid from the major UN aid agencies is listed in Table 24.5.

Transfers

24.3.3 The EU

Development assistance from the EU is organized by the European Investment Bank (EIB) and the European Development Fund (EDF), and is tied closely with trade policy. The assistance is provided according to framework agreements for the co-operation between the EU and 77 participating developing countries in Africa, the Caribbean and the Pacific. The Cotonou framework agreement (named after the capital of Benin where the agreement was signed) for the period from 2000 to 2005 is financed by EDF, with a budget of approximately US\$15.2 billion. In 2000, the net concessional disbursement by EU to countries on 'Part II' of the DAC List amounted to US\$4,414 million, and the non-concessional disbursement was US\$427 million.

24.4 Bilateral donors

Bilateral aid is transferred directly between national governments and recipient countries. Most developed countries have established national aid organizations either as integrated parts of their ministries of foreign affairs, as in Denmark, or as separate institutions, as in Sweden. Normally, field offices are established in those developing countries that receive extensive assistance. The offices supervise activities funded by

Organization

the donor countries and have a continuous dialogue with host countries about overall development strategies, sector strategies, identification of support areas and about monitoring of progress of on-going projects and programmes financed by the donor. Normally the field offices form part of the respective embassies or high commissions. The level of decentralization of decision-making varies greatly from donor to donor. For some years there has been a general trend to decentralize more decisions to field offices. Bilateral aid is offered in a variety of forms including grants, concessional loans and subsidies to export credits in the form of so-called mixed credits.

Grants Grants are gifts requiring no repayment. They are attractive to developing countries, as they do not directly increase the debt burden, which has already reached unsustainable macro-economic levels in many cases. Grants may be provided in many forms, such as general budget support for payment of debt service obligations to the different development banks, or as general support for government expenditure. Grants may also be used for support to a given sector, such as budget support for road maintenance. TA and TC are other support areas often financed by bilateral grants. There has been a growing level of bilateral general budget support as part of the move towards ‘sector-wide approaches’ (SWAP) or ‘sector programme support’ (SPS) in an effort to have a more ‘comprehensive development framework’ (CDF) – again based on ‘poverty reduction strategies’ (PRS). The concern of donors in this respect is that such an approach requires strong institutions for planning, implementation and cost control – a condition that is lacking in many developing countries.

Monitoring Because of such weaknesses, donors often wish to support specific projects or programme components where special monitoring systems and cost control systems can be set up. In this context, road construction projects are relatively easy to identify, specify, tender, execute and hand over. Although this is an area where corruption is widespread, it is possible for donors to insist on having qualified technical/financial audits to check if value-for-money has been achieved.

Soft loans Some donors provide concessional loans on a bilateral basis. An example is Germany through the Kreditanstalt für Wiederaufbau (KfW).

Mixed credits A number of bilateral donors are using part of their ODF to subsidize export credits. In order to minimize the potential negative effects on competition for projects that are commercially viable, the DAC countries have agreed a set of rules for the use of mixed credits. The main point is that only projects that are not commercially viable, but are good long-term development projects, can be supported by mixed credits. Wind farms are one typical example. Ports and airports may also fall into this category.

24.5 Assistance to the transport and road sector

Transport is an essential means for enhancing economic and social development. It is indispensable for the efficient development of most social and economic sectors and donor countries direct a substantial proportion of the development aid to the transport sector. In 2000, approximately 10 per cent of all bilateral transfers and 13 per cent of multilateral transfers were allocated to transport and communications. This amounted to more than US\$9 billion in 2000, and the distribution of the financing is shown

Table 24.6 ODF to transport and communications (2000)

	<i>Total disbursements (%)</i>
Total bilateral	9.6
Total multilateral	13.1
World Bank	14.7
Regional development banks	13.0
European Commission	9.4

Source: OECD (2002).

in Table 24.6. Bilateral support to the transport and communications sectors varies greatly from donor to donor. For example, Denmark and Japan provided about one-quarter of their assistance to the sector in 2000, but several donors provided less than 3 per cent.

24.6 Recent trends in aid management

In the early years of development co-operation, aid was provided in many different forms to developing countries. The interventions were often for relatively short-term projects in many different sectors, without much analysis of the relationship to the wider sector or national context. Donors often had different demands for involvement and financial contributions from the recipient institutions. Implementation was, in many cases, carried out by donor-financed advisers. Gradually, it became apparent that many expected project benefits were not achieved or sustained, because too little attention had been given to the wider context, such as macro-economic stability, institutional capability, lack of finance for maintenance, etc. The awareness of the need to ensure full ownership of the projects, institutional capacity and financial resources for sustained services increased during the 1990s. Also, the need for a more comprehensive and co-ordinated approach has been recognized. More recently, this recognition has led to an increasing number of bilateral and multilateral institutions to adopt PRS as a framework for development assistance. There is a growing awareness among institutions that this must be accompanied by a harmonization of donor procedures to reduce the administrative burden for the partner countries and the transaction costs. A number of implications are emerging from this and are discussed below.

In order to develop mutual trust and a better understanding of broader contexts, it is important that donors and recipient countries agree on long-term co-operation in key sectors in accordance with the PRS, with the recipient country taking the lead. Many small bilateral donors have decided to reduce the number of co-operation countries and sectors to strengthen partnerships. The choice of sectors is decided in close consultation with the recipient country.

Donor co-ordination is facilitated if national strategies (PRS and sector strategies) are the basis for all donor support. It is important that the recipient country has

Evolution

Long-term partnerships

Donor co-ordination

sufficient capacity to administer on-going improvements and adaptations of the national strategies, and the implementation and associated controls of agreed support activities. If this is the case, the donors should adopt common principles and procedures.

Sector programme support	As a result of these trends, many donors have moved from single project support to broader and long-term sector programme support within the context of national sector strategies and plans. Such programmes often comprise several major components and have sizable budgets for periods of about 5 years. The level of detail is often less than in traditional projects. Some interventions are detailed during the programme period. This requires flexibility and procedures for review and decision-making over time. The move towards sector programme support does not mean an end to the project concept. Within a programme, several projects may be defined at the outset or may emerge during the programme period. As an example, a transport programme may contain a well-defined major rehabilitation of a given stretch of road, which will be planned, tendered and executed along traditional lines.
General budget support	The ultimate ideal aim would be to provide donor funds to the general government budget of recipient countries without special ties to sectors, programmes or projects. This would require a fully acceptable PRS, that the public expenditure budget reflects the priorities set out in the PRS, that the recipient institutions at all levels have the capacity to implement the activities reflected in the budgets, and that satisfactory mechanisms are in place to monitor progress, report, control costs, and to undertake technical and financial audits. In most developing countries, these conditions are not fully met. Therefore, for a number of years, general budget support on a large scale is not likely. However, general budget support is granted in some cases, notably as part of macro-economic structural adjustment programmes, including debt relief programmes.
Sector budget support	In some sectors, budget support is relatively advanced. This is the case in health and education programmes in some countries where institutional reforms have been adopted, institutions strengthened and financial control mechanisms put in place. Much of the budget support to the social sectors covers recurrent cost. In an infrastructure sector, like roads, the investment budgets are relatively high and often supported by the donors, while it is expected that the recipient governments fund the major share of recurrent costs for operation and maintenance.
Role of technical assistance	The role of technical assistance/co-operation funded by the donors is also changing. In the past, this was often associated with planning and execution of specific donor projects. Now, broader assistance is provided to sector policy reforms, regulations and to institutional capacity building.
Aid management	An example of the current aid management practice by a bilateral donor is shown in Box 24.1
Involvement of the private sector	There has been a move by donors and developing countries towards involving the private sector more and encouraging the state to concentrate on its regulatory role (see Chapter 22). This is particularly the case in the telecommunications sector, but applies also to the energy and transport sectors. The private sector can play an important role in the execution of investment projects, in operation and provision of services, and in financing.

*Box 24.1 Aid management in Danida**Organization*

The Danish International Development Assistance (Danida) is a fully integrated part of the Danish Ministry of Foreign Affairs. The headquarters of Danida is in Copenhagen, and there are offices within Danish embassies in most countries receiving substantial Danish bilateral aid.

New approach to aid delivery

In the mid-1990s, Denmark adopted a new strategy aiming at deeper and longer term co-operation arrangements with fewer developing countries. Twenty programme countries were selected, mostly in Africa. A long-term co-operation strategy was agreed with each partner country defining 2–4 sectors that would receive the bulk of the Danish assistance. Instead of traditional project assistance, the aim was gradually to provide aid as SPS. Project assistance could still be provided to certain other sectors, notably for human rights, good governance, balance-of-payment support and general reforms in the public sector. The strategy was subsequently updated and called ‘Partnership 2000’ (Danida 2000). The same basic principles were maintained, but with increased emphasis on partnership and the principle that development aid should support the developing country’s own efforts – as mapped out in its PRS and its specific sector strategies and plans.

Programme countries

In 2002, the number of main programme countries was reduced to 15:

- Africa – Benin, Burkina Faso, Egypt, Ghana, Kenya, Mozambique, Tanzania, Uganda and Zambia
- Asia – Bangladesh, Bhutan, Nepal and Vietnam
- Latin America – Bolivia and Nicaragua.

Sector concentration

Within the programme countries, different sectors have been selected in close consultation with the partners. A total of 42 Danish-supported SPSs operate in sectors such as agriculture, health, education, water and sanitation, energy, transport and in a few cross-cutting ‘sectors’ including private sector development, environment and local government. Several of the early SPSs are already in a second five-year phase. In the transport/road sector, Danida has active SPSs in Bangladesh, Benin, Ghana, Nicaragua, Tanzania, Uganda and Zambia.

Sector policies

In order to create clarity and consistency in the approach to support in the various countries, Danida has formulated policies for support in the main sectors supported. One example of this is Danida’s transport infrastructure sector policy (Danida 1999).

Mixed credit

Danish bilateral aid is also supporting mixed credit projects, often in countries outside of the SPS programme. China, Thailand and several Latin American countries receive much of the mixed credit assistance mainly for the energy, water and sanitation, environment and health sectors. In the transport sector, mixed credit finance has been used for the improvement of airports, ports and railways.

NGOs

Danish NGOs also receive a substantial share of the bilateral aid. Much of this is used to finance humanitarian and social projects, but rural development projects are also common. These may involve improvements of road infrastructure, often using labour-based methods.

Decentralization of management

For the partnership to deepen, it is important that Danida is present with sufficient capacity and sector knowledge in the partner countries. This has been achieved to some extent by establishing Danish embassies in all programme countries. There are plans to strengthen the embassies further by posting more staff with sector knowledge and delegating more responsibility.

Aid management guidelines

In order to facilitate the decentralization, an updated and streamlined set of aid management guidelines is being issued. The guidelines can be accessed through the Danida homepage of the Danish Ministry of Foreign Affairs' website. The guidelines emphasize procedures and principles based on ownership by partner institutions of the processes and interventions, dialogue, and of donor co-ordination. There will be a move towards joint financing and budget support as soon as the partner institutions are capable of managing implementation, operation, sustainable maintenance, financial control and auditing.

Adapted from: Danida (2002).

24.7 Project cycle and framework

24.7.1 The concept

National
sector
framework

Most developing countries have an established set of policies, strategies and programmes, which may include detailed plans for interventions or activities for which they require donor assistance. These contain normally a substantial number of interventions, which have been subject to analysis of widely varying scope and quality. Governments in developing countries and donor agencies therefore deal with a large number of activities related to development and investments. These can be at various stages of identification, preparation and implementation. These activities are initiated by and under the overall responsibility of the government in a recipient country.

However, where external assistance is sought, the process involves close co-operation between the authorities of the recipient country and the donor agency, and also with private parties. A clear understanding of the elements involved in project work and the definition of the various phases is a basic requirement for effective co-operation. There is a need for a comprehensive and systematic framework for planning and administration.

The World Bank has played an important role in developing appropriate project planning and implementation tools. One such tool is the ‘project cycle’ concept, which was developed several years ago (Baum 1982). The concept is still useful in many situations in spite of the move from project approach towards the broader sector programme support approach, discussed earlier. It can be used for programme planning in an adapted form. The project cycle categorizes the various phases of a development assistance activity, from when it is first conceived as an idea through to its completion and use. This is done by application of a terminology, which is generally understood and accepted by donor agencies, the co-operating partners in the developing countries, as well as by private firms and institutions participating in the process. The terminology used for the phases differs between organizations (see, e.g. the terminology used in Chapter 7). Even the number of phases may differ. A clear line may not always be drawn between phases and some phases may contain elements of a preceding phase or the following one. Technically, the ‘project cycle’ refers to a sequence of projects, where experience gained from one is fed back into the next. However, the term is often applied to individual projects that have defined start and end points.

Terminology
and use

The project cycle can be divided into four main phases, often termed

Phases

- Identification
- Preparation
- Implementation
- Completion.

Each main phase is sub-divided into a number of sub-phases as illustrated in Figure 24.3. Engineering inputs may be involved in each of these phases, but there is likely to be significant engineering design work for a road project during the ‘preparation’ phase and for supervision during the ‘implementation’ phase.

24.7.2 Identification

The identification process is normally based on a request for support from a recipient country to a donor organization. It is undertaken as a joint effort between the partners and normally includes the elements described in Box 24.2.

24.7.3 Preparation

The government of the recipient country is normally responsible for preparation of a supporting feasibility study, but will in most cases receive assistance from the donor. The feasibility study is an in-depth analysis undertaken to establish whether the proposed project/programme will solve the problems raised during identification.

Feasibility
study

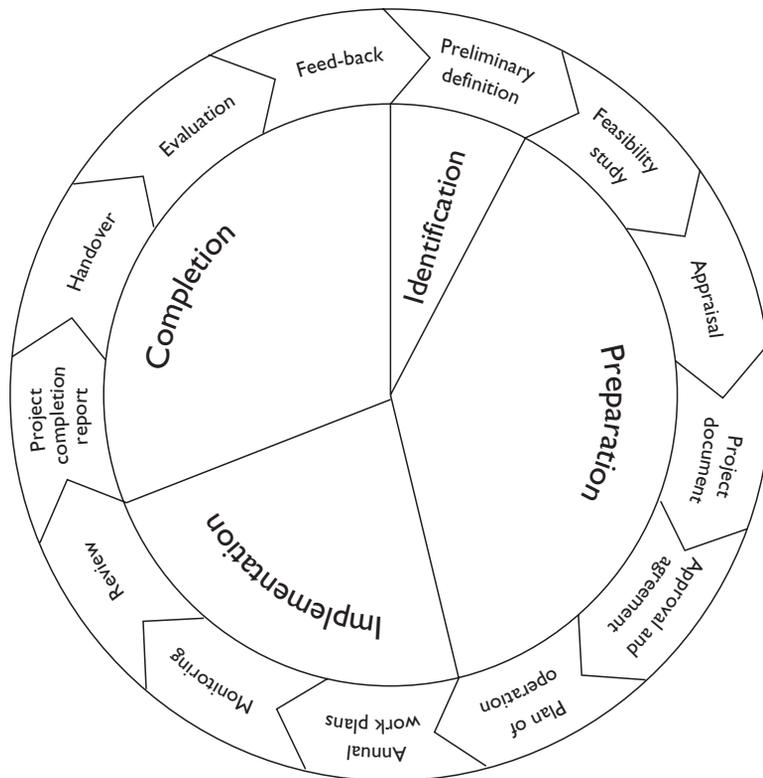


Figure 24.3 The project cycle.

Box 24.2 Identification

Both authorities of the recipient country and the donor agency are involved in analysing development strategies, such as PRS, national sector strategies and a country assistance framework, and in identifying interventions that support those strategies.

The identification process may involve sector analyses and studies, and these should be broad and inter disciplinary. As far as possible, these studies should build on previously conducted analyses and existing data. The need for additional analyses and studies will be assessed, not least to establish relevant indicators for monitoring of impact on poverty, gender, environment, human rights, good governance and democratization/participation. Also, additional data to assess macro-economic and institutional aspects may be needed. Duplication of effort should be avoided by ensuring co-ordination with other donor activities in the relevant sector.

If projects or programmes emerging from this analysis are evaluated to be financially, economically, socially and environmentally sound, and in conformity with both the country's and the donor's objectives, the preparation phase is initiated.

The feasibility study also determines whether the proposed project has the appropriate scope and the relevant inputs, whether it is possible to implement, to sustain in the future after the donor involvement has been completed, and whether alternative solutions should be considered. Provided that the proposal is acceptable, the feasibility study will conclude with the preparation of a draft project/programme document.

The appraisal is the donor's assessment of the proposed project or programme following the feasibility study. It is the last activity prior to submission to the authorities for funding and approval. The purpose of the appraisal is to provide decision-makers with appropriate information and documentation to enable rational decisions, and to contribute to a well-structured and realistic final project/programme document. Appraisal is solely the donor's responsibility. It is conducted by the donor's own staff, often supplemented by individual consultants. The appraisal report sets out its findings, conclusions and recommendations in clear, comprehensive and unequivocal terms as a basis for the ensuing decision-making process. Box 24.3 provides further background on the main elements of the preparation process.

Appraisal

Box 24.3 Preparation

Feasibility study

The feasibility study analyses the proposal from technical, economic, financial, social, institutional, organizational and environmental angles. The feasibility study typically requires a multi-disciplinary team as it addresses wide-ranging issues that are often inter-linked and complex. The study is normally undertaken by external consultants using terms of reference agreed between the recipient and the donor. The consultants recommend under which conditions the project is feasible or not. These recommendations are not binding for the donor or the recipient.

Feasibility studies must identify and prepare preliminary designs of technical and institutional alternatives and compare their respective costs and benefits. Most developing countries are characterized by abundant and low-cost labour and scarce capital. A donor will, therefore, look for technological solutions most appropriate for the country's resource endowment at its stage of development.

If required, an environmental impact assessment is undertaken as part of the feasibility study.

Appraisal

The appraisal provides the final check before approval of the project/programme document. The appraisal will focus on the degree to which the planned project/programme has the desired features, and which relates to both the country's and the donor's objectives. The appraisal will assess the degree to which these features are addressed in the national sector context. In addition to these over-riding issues, the document under appraisal must reflect the fact that the necessary technical, social, financial and economic, environmental and institutional analyses are adequately prepared. The appraisal should confirm that the

available documents (identification report, feasibility report, detailed engineering design, project/programme document) are the following:

- *Comprehensive* – all relevant aspects have been analysed based on sufficient data, with specific confirmation that assumptions and risks have been analysed sufficiently, and relevant mitigating measures included.
- *Consistent* – conclusions are supported by sound arguments.
- *Technically and financially valid* – solutions proposed and activities recommended are practical, cost-effective and realistic; and that an effective process and impact monitoring system is included.
- *In concurrence with national and donor policy concerns* – both are explicitly described and adequately addressed.

Approval and agreement

Different donors have different procedures leading up to final approval and provision of funds. However, all are based on an appraisal supported by specialized documents, such as a project document, and supplemented by customized presentation formats. Provided that the funding for the project is approved, the next step is to establish a government agreement for the project, as ODA is based on government-to-government co-operation. Normally, agreements are between a donor and the recipient's Ministry of Finance.

Plan of operation

In order to ensure a systematic and efficient execution of the project/programme document, a 'plan of operation' is developed, which serves as a management tool. This plan contains detailed descriptions of activities and divisions of responsibilities between the various parties involved, work schedules, budgets and the like.

24.7.4 Implementation

Responsibility

The responsibility for implementation will normally be entrusted to an agency of the recipient government. In some cases, technical assistance will be provided to the agency through advisers or consultants. The responsibility for procurement of works, goods and services will, in some cases, rest with the agency and, in others, remain with the donor – depending on the donor's procurement regulations.

Annual plans

In order to ensure the timely implementation of the project, detailed annual work plans are normally prepared based upon the plan of operation. In principle, the implementing agency is responsible for the preparation of the annual work plans based on inputs from consultants and contractors. The annual work plan provides the detailed implementation information, such as activity diagrams, budgets, etc., and constitutes a framework for making revisions to the original overall plan of operation in the case of unforeseen difficulties.

Donor supervision

Supervision takes place in a variety of ways. During negotiation, agreement will have been reached on the schedule of progress reports to be prepared by the implementing agency. These may cover the execution of the project, its costs, the financial status of revenue-earning enterprises and information on the evolution of project benefits. Box 24.4 shows some of the instruments used for supervision of implementation.

Box 24.4 Tools used to assist supervision of implementation*Progress reports*

Progress reports and site visits form the basis for the donor's project supervision and utilize the monitoring indicators established at the time of appraisal.

Site supervision

Site supervision by the donor is often placed in the hands of consulting firms, and progress reports are reviewed at the resident mission. Problems are resolved locally or at head office, as appropriate.

Monitoring

The purpose of monitoring is to provide the necessary information for project management, which the recipient and the donor need to adjust activities, inputs and budgets to ensure achievement of objectives.

Periodic reviews

During implementation, the donor undertakes periodic reviews, often jointly with the recipient and/or with other donors. A review is a comprehensive assessment of the implementation compared to the project/programme document, the government agreement and the annual implementation plans. The purpose of the review is to identify deviations and problems, to adjust the implementation plans accordingly, and to set out clear guidelines for the project's development until the next review.

24.7.5 Completion

Once a project is completed, a project completion report is prepared by the donor. The report provides a systematic description and evaluation of the preparation, implementation and outcomes of the completed project to document the activities of the project cycle and to extract valuable experience. Normally the project completion report is prepared without field visits and is based on existing documents.

A formal evaluation is carried out for many projects. This is a systematic and objective assessment of preparation, implementation and project outcomes. The purpose of an evaluation is to assess the appropriateness of the objectives, the rate of achievement of objectives, effectiveness and sustainability. Most development assistance evaluations are carried out 'ex post', that is some time after the project is completed. Most donor agencies have established autonomous evaluation units and utilize independent consultants to carry out the evaluations. These are undertaken through a combination of desk and field studies, which may involve gathering primary field data. In World Bank evaluations, it is customary that the recipient is requested to prepare a separate section with an evaluation of the Bank's performance.

The results of an evaluation must be highly reliable and presented in a form that enables the findings to be incorporated in the donor and recipient's future decision processes in connection with new projects. The evaluation report is the most important single feedback document in development work. Through the evaluation, lessons

Completion
report

Evaluation

Feedback

learned are fed back to the parties involved, and built into the design, preparation and implementation of future projects. Hereby the project cycle is completed.

References

- Baum, W.C. (1982) *The Project Cycle*, Washington, DC: The World Bank.
- Danida (1999) *Transport Infrastructure: Danida Sector Policies 1999*, Copenhagen: Ministry of Foreign Affairs, Danida.
- Danida (2000) *Denmark's Development Policy-strategy: Partnership 2000*, Copenhagen: Ministry of Foreign Affairs, Danida.
- Danida (2002) 'Denmark's development assistance', *Annual Report 2001*, Copenhagen: Ministry of Foreign Affairs, Danida.
- OECD (2002) *Development Co-operation 2001 Report: Efforts and Policies of the Members of the Development Assistance Committee*, Paris: Organisation for Economic Co-operation and Development.



Going home from school in Vietnam. (Photo: courtesy TRL Ltd.)

Index

- AAADT *see* Annual average daily traffic
AASHTO Road Test 287, 289
AASHTO *see* American Association of State Highway and Transportation Officials: pavement design method 285, 289–94, 303
Abrasion 164
Abutment 78, 204
Acceleration 232
Access 12, 15–17, 19, 20, 36, 67, 85, 86, 96, 97, 102, 110, 111, 127, 131, 206, 374; basic 127; charge 445; road 208, 210, 215, 219, 230
Accessibility planning 111
Accident: costs 59–62, 372; traffic *Chapter 4*, 38, 80, 117, 178, 206, 222, 229, 239, 442
Accounts/accounting 32, 442, 447, 457
Accuracy 118
Action plan 71, 75
Activities 372
Additives 267
Adjudication 341
Administrator/administration 28, 452–6
Adoption, road 36
Advance payment 333, 335
African Development Bank/Fund (AfDB/AfDF) 487
African socialism 7
Agency arrangements 310, 453, 456
AGETIP 310, 311
Agglomerate 164
Aggregate 154–7, 163, 164, 167, 177, 257, 264, 273–5; loss 398
Aggregate crushing value (ACV) 156, 268
Aggregate impact value (AIV) 156, 254, 255
Agriculture 14, 17, 18, 20, 79, 82, 85, 106, 111, 121, 131, 173, 175, 237, 347, 483
Aid/development assistance *Chapter 24*
Aid management 491–4
AIDS 10, 489
Air/aerial photographs 138–41, 185, 237, 239
Air quality 82, 83, 87, 89
Alcohol 64
Alignment *Chapter 12*, 80, 82, 84, 137, 193, 208, 257, 375
Allophane 175
Alluvial 139, 176, 177
American Association of State Highway and Transportation Officials (AASHTO) 147, 151, 154, 157, 205, 214, 248, 266, 330
American Society for Testing Materials (ASTM) 147, 330
Analysis period 115, 116, 123
Analytical pavement design 285, 295–302
Anionic emulsion 267
Annual average daily traffic (AADT) 42, 46, 48, 50, 53–5, 215, 216
Application rate, binder, chippings 270–2
Apron, culvert 202, 203, 412
Arbitration 315, 326
Archaeology 84
Area, rainfall 180
Area reduction factor (ARF) 180
Armco 195
Arterial 208, 210, 216, 230, 374
Asbestos 82
Asgronden 175
Ash 164
Asian Development Bank/Fund (AsDB/AsDF) 64, 487
Asphalt *Chapter 14*, 168, 171, 173, 350, 395–402; cement 264; mixes 278–82; paving 282, 283
Asphaltic concrete 177, 273–8
Asset 24, 36, 107, 371, 372, 450
Assistance to transport and road sector 490, 491
ASTM *see* American Society for Testing Materials
ATTAC 9

- Audit: financial and technical 389, 447, 457; safety 67–70, 76
 Auger 146
 Australia 63, 75, 167, 170, 171
 Austroads 301
 AutoCad 242
 Availability fee 449
 Average least dimension (ALD) 268–71
 Axle load 38, 43, 53, 55, 56, 124, 287, 294, 296, 376
 Axles, equivalent standard (ESA) 43, 286–90, 294, 302

 Back-analysis, pavement design 297
 Baghouse fines 275
 Bangladesh 62
 Bank guarantee 333
 Bar chart 336, 337
 Basalt 163, 172
 Base course *see* Binder course
 Baseline 114
 Base, road 252, 254–6
 Batching plants, asphalt 83, 279–81
 Bedrock 138, 139
 Behaviour, road user 63, 64, 73, 74, 208
 Beijing 41, 43, 51
 Benchmark 32
 Benefits, benefit estimation 119–22, 417, 452, 453, 457, 458, 491
 Benkelman beam 303, 304
 Benzene 83
 Bhutan 245
 Bicycle 17, 18, 41, 44, 50, 127, 205
 Bilateral assistance/donors 484, 489–91
 Bill of quantities 118, 315, 322, 332, 339
 Binder, bituminous 264–8
 Binder course 252
 BISAR 296
 Bitumen 264–8, 355; affinity 156, 164, 165, 171
 Black cotton, black turfs 171
 Black market 123
 Blackspot 71
 BOO/BOT/BOOT/BOOST
 build/own/operate/subsidize/transfer 312
 Booze bus 75
 Borehole 146
 Borrow, borrow pit 81, 82, 84, 137, 242, 247
 Box culvert 196
 Braking distance/time 212
 Brandt Commission 8
 Brazil 166
 Break, emulsion 266, 269
 Breath test 64
 Bridge 78, 137, 171, 189, 195, 196, 202–4, 216, 225, 247, 327, 353
 British Standard (BS) 151, 266, 267, 330
 Brooming 273, 282
 Budget, budgeting 14, 24, 36, 93, 106–9, 126, 127, 352, 419, 430, 492
 Bump integrator 378–81
 Burning 408
 Bus 73; lane/stop/busway 51
 By-pass 85, 89

 CAD/E *see* Computer-aided design/engineering
 Calcrete 165
 Calibration 281, 296, 378–80, 427
 California bearing ratio (CBR) 155, 173–5, 253, 255, 256, 261, 287–9, 294
 Camber 356, 357
 Cambodia 167, 489
 Cameroon 64, 175
 Canada 448
 Capacity: implementation 107; traffic 14, 38, 41, 43–51, 207
 Capital: funds 443, 444; intensive 345; internationalization of 8; shortage 7
 Capping layer 253
 Carbonation 261, 262
 Carbon, carbon dioxide, monoxide 82, 83
 Career development 354
 Carriageway 216–19; marking 50
 Cart 50, 205, 358, 360
 Cartel 123
 Cascade 411
 Cash-flow 118, 353
 Casualty 58
 Catchment 179, 180, 184, 188, 239
 Cationic emulsion 267, 269
 Causeway 204, 413, 414
 Cement 168; modification 258; stabilization 257–9
 Certificate, payment 335, 336
 Champion 34
 Channelization 50, 78
 Channel modification 78
 Check dam 79
 Chief resident engineer (CRE) 325, 327
 Child 59, 72, 76
 Chile 45, 67
 China 17, 58, 355
 Chippings 268–73
 Chute 191, 192
 Cill 204
 Cinder 164
 Claim 337–41
 Class: environmental 86; road 36, 67, 76, 93, 107, 208–11, 215, 346, 374, 419
 Classification, traffic 39
 Classroom 463

- Clay 167, 169, 173, 175, 257, 259
- Clearing: cleaning drainage structures
410–12; site 244
- Client 36, 69, 450, 452–4
- Climate 81, 124, 172, 190, 346, 347
- Climbing lane 231
- Clothoid 228, 229
- Coastal plain 176
- Coefficient of uniformity 148
- Cold-mixed asphalt 355
- Cold war 8, 483
- Collector 208, 210, 216, 230, 374
- Collision diagram 71
- Colloid mill 267
- Colombo 58
- Comfort criteria 234
- Commercialization 24, 36, 443, 456, 457
- Commitment, project 115, 390
- Communication 479, 480
- Communism 9, 10
- Community 110, 128, 448; participation
19, 36, 111, 352
- Compaction 79, 151–3, 168, 169, 171, 175,
242, 249–51, 253, 255, 259, 261, 262,
282, 331, 355
- Compensation 86
- Competition 36, 310, 311, 442, 451–4
- Completion, project 499
- Complex modulus, bitumen 264
- Compound ranking 128–30
- Comprehensive development framework
(CDF) 490
- Computer-aided design/engineering
(CAD/E) 242, 243
- Computer, computerize 109, 126, 242,
243, 416
- Concession 312, 313, 449, 453
- Concrete 171, 195, 196, 202, 203, 328,
330, 417
- Condition, road 93, 95, 111, 382, 383
- Conditions of contract 316, 322
- Congestion, traffic 14, 41, 46, 83, 85, 93,
284, 417
- Consequence analysis 386
- Consolidation 248
- Construction, stabilized 244, 261–3
- Consultant 34, 35, 90, 325, 473–5
- Consultation 31, 35, 88, 89, 442, 457
- Consumer: consumption 106, 121; surplus
128, 131
- Contamination 81
- Contract: contractor *Chapters 16 and 17*, 34,
84, 89, 352–4, 386, 415, 447, 452, 453,
456, 458; documents 321–3, 326, 335
- Contractor development 367
- Contributory factors, accidents 62–4
- Control point 235, 236
- Copenhagen 44
- Core business 451
- Corporatize, corporatization 450
- Corridor 96, 99, 119, 137, 236
- Corrosion 195, 412
- Corrugation 403
- Cost 179, 312, 332, 352, 364, 366, 372,
386, 416, 417, 427–30, 442, 448, 449;
capital/investment/construction 117, 121,
189, 229, 230, 237, 246; control 366;
estimating 117–21; operational 189;
present value (PVC) 123; recovery
444–6; recurrent 117; standing 121
- Cost–benefit analysis 36, 60, 88, 467–70;
major roads 122–7; minor roads 128, 131
- Cost-effective 24, 81, 82, 85, 128, 129, 345,
347, 348, 378
- Cost-plus 313
- Cost-reimbursable contract 315
- Cost-sharing 36, 448
- Cotonou 489
- Countermeasures 70
- Counterpart training 463, 464
- Course description/notes, training 473, 474,
477, 478
- Court award 61
- Crack 173, 258, 261; repair, sealing
396, 412
- Crew, vehicle 120, 121
- Cross drainage 193–6
- Crossfall 190, 220
- Crossroad 239
- Cross-section 50, 80, 186, 197, 202,
215–22, 224, 242, 243, 248, 375, 410
- Cuba 166
- Cultural heritage 84, 85
- Culvert 78, 137, 171, 189, 190, 195–9, 216,
237, 247, 327, 353, 411–13
- Cure 261, 262, 266
- Currency 350
- Curvature 67, 78, 220, 221, 225–9,
231–4, 375
- Curve length 227
- Curve widening 219
- Customer 38, 452, 457
- Cutback bitumen 266, 267, 269, 282, 355
- Cut-off wall 195
- Cutting 78, 137, 162, 171, 222, 244, 245,
250
- Cyclist 24, 41, 42
- Dam 81
- Data 116, 117, 131, 185, 421, 422; accident
66, 71, 75; collection processing 53, 377;
traffic 38, 51–6

- Database 421, 422
 Dayworks 322, 334
 DBFO *see* Design, build, finance and operate
 Death *see* Fatality
 Decentralization 36, 110, 128, 352, 490, 494
 Default, contractor 341, 342
 Defect, road 376, 377
 Defects liability period 336
 Deflection, bowl 302, 303, 380
 Deformation 274, 380
 Delay 339
 Delegation 456
 Delta 139, 176
 Demand curve 122
 Demand, traffic 41, 95, 117, 448
 Demography 102
 Demonstration project 350
 Dendritic drainage 139, 140
 Denmark 485, 489, 491
 Density 150–2, 155, 171, 249, 250, 255, 258, 261, 276, 295, 331
 Dependency theory 6
 Depreciation 120, 121, 366
 Deregulation 9
 Desert 156, 170, 171, 298
 Design: and build 311; catalogue, pavement 288, 289; class/speed, geometric 209–11, 213, 222, 228, 230, 232; costs 119; discharge, culvert 196; life 286; project 115
 Designation, road 36
 Design, build, finance and operate (DBFO) 312
 Deterioration 401, 402, 421, 424–7, 442
 Development *Chapter 1, Chapter 24*, 92, 103, 206, 372
 Development Assistance Committee (DAC) 12, 484–6, 489, 490
 Development theory, strategy 3, 6–9
 Digital terrain model (DTM) 242
 Direct labour *see* In-house procurement
 Disability-adjusted life years (DALYs) 57
 Disabled 33, 57
 Discharge 186, 199
 Discounting, discount rate 123, 125, 126, 416
 Dispute 326
 Dissemination 31
 Distributor, bitumen 282
 Ditch 191, 199–202, 243, 244, 354, 411; check 201; intercepting 191, 192
 Divestiture 450
 Dividend 123
 Document, policy 31, 35
 Dolerite 163
 Domestic resources 346, 350
 Donor 7, 14, 34, 484, 490, 491, 495, 498
 Do-nothing, do-minimum 114, 120, 125, 418
 Dragging 403–5
 Drainage *Chapter 10*, 79, 139, 174, 225, 236, 237, 239, 242, 245, 257, 302, 409–14
 Drawing 323, 325, 326
 Drift 170, 171, 194, 413, 414
 Drink driving 75, 76
 Drive cylinder 150
 Driver 51
 Drop structure 202, 204
 Drum-mixer 281, 282
 Dry-bound macadam 256
 Dry compaction 249
 Dualism 6
 Dune 139, 170
 Durability, asphalt 274
 Dust 83, 164, 281, 376
 Dynamic cone penetrometer (DCP) 154, 288, 302, 381
 Earthworks 225, 242, 244–53, 350, 354, 355
 East African Flood Model, TRRL 181
 East Timor 489
 Ecology 10, 79, 80
 Economic appraisal, analysis *Chapter 7, Chapter 21*
 Economic development, and benefits 14, 15, 107, 121, 122, 206, 346, 458
 Economic growth 5, 14, 16, 92, 106, 109
 Economic imperialism 6
 Economic reform 9
 Economic return 114
 Economic stability, macro 491
 Ecosystem 80, 83
 Edge deterioration 397, 398
 Effectiveness 441–3, 451, 452, 455, 457, 498
 Efficiency 106, 126, 207, 312, 371, 441–3, 451, 452, 455, 457, 458
 Egypt 74
 Elasticity 100–3, 122
 Elastic modulus (E-modulus) 296–301
 Electrical resistivity 145
 ELSYM5 296
 Embankment 84, 137, 162, 175, 195, 202, 247–50
 Emissions 81, 83, 87, 429, 430
 Empirical pavement design/model 285–94, 425
 Employer 316, 325
 Employment 10, 92, 106, 116, 121
 Emulsion, bitumen 257, 267, 269, 271, 282, 355

- End-product specification 154, 250, 314, 329, 330
Endwall 195
Energy 10, 198, 429
Enforcement 63, 64, 73, 75, 76
Engineer 316, 325–7, 361
Engineer, procure, construct (EPC) contract 311, 312
Environment *Chapter 5*, 9, 24, 36, 46, 63, 93, 109, 117, 124, 208, 247, 285, 286, 293, 372, 417, 419, 421, 442; training 470, 471
Environmental impact assessment (EIA) 78, 86–90
Enzyme 257
Equipment 14, 83, 84, 246, 249, 250, 345, 346, 347–50, 356–60, 364–7, 405, 406
Equity 106, 109
Equivalence factor (EF) 43
Erosion 78, 79, 81, 82, 139, 140, 163, 164, 170, 178, 179, 186, 192, 199–204, 222, 236, 242, 403, 409–12
Ethiopia 17
Euler spiral 228
European Bank for Reconstruction and Development (EBRD) 312, 487
European Commission, European Union (EU) 9, 86, 90, 445, 487, 489
European Development Fund (EDF) 489
European Investment Bank (EIB) 489
Evaluation of tenders 319, 320
Evaluation, project/course evaluation 85, 110, 115, 477, 498
Evaporation, evapo-transpiration 171, 288
Excavation 244–7
Exchange rate 123, 365
Expansive clay 171–5, 249
Expression of interest 317
External training 463, 464
Extreme value statistical analysis 181, 183
Extrusive rock 163
- Failure, pavement 43, 284, 296, 297, 301, 302
Falling weight deflectometer (FWD) 297, 298, 302, 303
Fan, alluvial 176, 177
Farm, farmer 89, 121
Fatality 57–60, 72, 75
Fatigue, pavement 298–301
Feasibility, study 93, 109, 115, 495
Ferricrete 166
FIDIC 315–23, 325, 330–3, 336–41
Field investigation 144–7
Fill 78, 162, 174, 248
Filler 275
Filling depressions 396, 397
Filter 254
Finance 93, 95, 106, 107, 312, 313, 441–9, 452, 456, 491
Finland 448
Fish 79
Flakiness index 156
Flood discharge estimation 183–8, 190
Flood, plain 78–80, 139, 166, 176–8, 192, 202, 203, 237
Flora and fauna 79–81, 85
Flow: bitumen 277; traffic 38, 42, 43, 51, 93
Fog spray 399
Foliation 165
Food security 10
Food for work 488
Force account *see* In-house procurement
Ford 193, 194, 413, 414
Forecasting 32, 38, 54, 55, 98–106, 115, 117, 376, 419
Foreign currency, exchange, debt 7, 371
Formation 252
Forms of contract 313
Framework, policy 26, 27, 29, 35
Free swell 149, 150, 173
Free trade 9
Friction 208, 212, 213, 226, 227
Fuel 120, 207; levy 14, 445–7
Functional classification 208
Functional separation 450, 452, 454, 458
Functional specification 314
- Gabion 203
Gang leader *see* Supervisor
Gap-graded 168
Gazetteer 36, 375
Gender 18, 19
Geographic information system (GIS) 87, 96
Geometric design *Chapters 11 and 12*, 67, 76, 96
Geophysical exploration 144–7
Geotechnical costs 118–19
Ghana 6, 17
Ghandi 6
Ghost island 240
Global cost 118, 128, 242
Global development theory 8
Globalization 9
Global Road Safety Partnership (GRSP) 65
Gneiss 165
Goods transport 106, 111
Government, governance 9, 26, 28, 34; policy 24–6; transfers 448
Gradient 120, 191, 207, 229–32, 240, 241, 375

- Grading *see also* Particle size distribution 403
- Granite 163
- Grant 490
- Grass 79
- Gravel 167, 176, 254–7, 355, 376
- Gravity model 102, 103
- Green paper 31, 34, 35
- Green revolution 7
- Grip 191
- Gross domestic product (GDP) 4
- Gross national income (GNI) 3–6, 62, 75, 100, 484, 487
- Gross output 61
- Growth factor model 101–4
- Growth, traffic 38, 42, 98, 99, 102, 106, 376
- Grubbing 244
- Guidepost 414
- Gutter 191
- Guyana 176
- Gyratory compactor 275
- Habitat 78–80
- Halloysite 175
- Hanoi 58
- Harare 58
- Hardpan 167
- Harness, animal 360
- Hawaii 166, 175
- Head-load, basket 100, 357, 358
- Headwall 195
- Headwater 196, 197
- Headway 44
- Health 64, 83, 84, 353
- Helmet, crash 72, 76
- Herbicide 408
- Hierarchy, road *see* Class
- Highway development and management tool (HDM-4) *Chapter 21*, 45, 93, 109, 125, 294
- Histogram 33
- HIV 10
- Horizontal alignment 224–9, 242, 243
- Hospital 58
- Hot-mix 273–5, 355
- Hot-rolled asphalt 273
- Human capital 61, 62
- Human development/poverty index (HDI/HPI) 5, 6, 11
- Human resource development 462
- Human rights 9
- Hydraulic design 188–90
- Hydration 258, 261, 262
- Hydrocarbons (HC) 83, 264
- Hydrograph 184
- Hydrological year 182
- Hydrology *Chapter 10*
- ICE *see* Institution of Civil Engineers 315, 332
- Identification, project 115, 495
- IGN *see* Institut Geographique National
- Igneous rocks 163, 164
- IKONOS 141
- ILO *see* International Labour Organization
- Impact, environmental 78–86
- Implementation: project 498, 499; training 479–82; works 115, 386, 389, 414, 415
- Import: control/quota 14, 36; substitution 7
- Improvement, low-cost 70, 76
- Income 4, 12, 102, 106; per capita 4, 103
- Incremental analysis/BCR 96, 431
- India 6, 7, 11, 64, 67, 167, 171, 172, 355
- Indonesia 166, 171, 175, 264, 356
- Industrialization 8
- Inflation 123, 365, 447
- Information: management 88, 95, 374–6, 442; traffic 51–6
- Information quality level (IQL) 52, 53, 421
- Information system: finance, accounts, personnel 457; road maintenance 390–2, 457; road safety 64–6, 75
- In-house procurement 309, 310, 352, 353, 386, 414, 415
- Injury 57
- Inlet 195, 202; control 197, 198
- InRoads 242
- Inspection 329, 377–9
- Instability, slope 245
- Institut Geographique National (IGN) 239
- Institutional arrangements 110, 116, 352, 462, 463, 491
- Institutional change/development *Chapter 22*, 35
- Institution of Civil Engineers (ICE) 315, 332
- Instructions: to tenderers 320; training 481
- Instructor, driver 73
- Instrument, policy 26
- Integrated policy 28–31
- Integrated rural accessibility planning (IRAP) 112
- Integrated rural development programme 14
- Intensity-duration curve 184, 187
- Inter-American Development Bank (IADB) 487
- Interceptor ditch 79, 82
- Interest: rate 123, 335
- Intermediate technology 345
- Intermodal 95
- Internal training 463

- International aid agency 487–9
 International Bank for Reconstruction and Development (IBRD) 487
 International competitive bidding (ICB) 310, 320
 International Development Association (IDA) 13, 487
 International Finance Corporation (IFC) 487
 International Fund for Agricultural Development (IFAD) 487
 International Labour Organization (ILO) 364, 488
 International Monetary Fund (IMF) 7, 483, 487
 International Road Maintenance Handbook *Chapter 20*
 International Roughness Index (IRI) *see* Roughness
 Intersection 42, 50, 67, 239–41; sight distance 214, 215
 Intrusive rock 163
 Inventory 36, 95, 375
 Invert 412
 Investment *Chapter 21, 9*
 Iran 170
 Iraq 176
 Ironstone 166
 Island, traffic 240
 Isohyet 180
 Ivory Coast 487
- Jakarta 45
 Jamaica 64, 67
 Japan 9, 175, 485, 491
 Job description 467, 468
 Joint services committee 454
 Joint stock company 454
 Junction, road 67, 206
- Kaolin 166, 169
 Karachi 58
 Kenya 7, 13, 67, 355
 Kerb 240
 Kirpich formula 187
 Knowledge, road user 63, 64
 Kosovo 489
 Kreditanstalt fuer Wiederaufbau (KfW) 490
 K-vlaue 232
- Laboratory testing, soils 147–57, 328, 329
 Labour-based methods 14, 18, 222, 345–64, 403, 406, 488
 Labour, international division of 8
 Lacustrine 176
 Land facet, system 142
- Landsat 141
 Landscape 78, 176
 Landslide *see* Slip
 Land use 67, 75, 179, 186, 187, 239
 Laterite 166–9, 255
 Law *see* Legislation
 Lead 82
 Legislation 23, 26, 31, 35, 36, 72, 75, 338, 354, 442, 458
 Lesson plan 476
 Level-of-service 24, 29, 36, 46, 47, 93, 106, 107, 109, 127, 237, 372, 374, 443
 Licence 446, 447
 Life cycle analysis/costing 93, 124–5, 257, 284, 285, 416–18, 423, 424
 Life insurance 61
 Lighting 33, 70
 Lime 168, 175, 253–5; modification 260; stabilization 259–63, 355
 Limestone 162, 164, 165, 172, 257, 259
 Linear elastic theory 295, 296
 Line profiling 145
 Linear programming 106
 Lining, ditch 200, 201
 Link 422
 Liquidated damages 335
 Liquid limit (LL) 149, 157, 173, 253
 Local competitive bidding (LCB) 310
 Local sealing 395
 Logit model 103
 Longitudinal drainage 190–3
 Los angeles abrasion 156, 157, 169, 254, 255, 330
 Low-cost road alignment 241, 242
 Lump sum 313
- Macadam: dry-vibrated 256; penetration 355, 356
 Magma 163
 Maintenance *Chapters 19 and 20, 207, 427, 428*
 Major road 114, 115, 443, 446
 Malawi 62
 Management 21, 23, 35, 36, 38, 52, 84, 364, 450–7; contract 315; cycle 373, 374; functions 373; maintenance, road network *Chapter 19*; system 109
 Manager, network 453, 454
 Manhole 414
 Manning 186, 192, 193, 197
 Mantle rock 166
 Manual, worker 23, 471
 Map, mapping 84, 96, 111, 137, 138, 179, 185, 237, 241
 Marble 165
 Margalitic 171

- Marginal materials 255
 Market economy 9, 14, 20–1, 444, 445, 455
 Market price 122, 123
 Marking, road 70, 220, 327
 Marshall method 275–9
 Marsh soil 176
 Mass-haul curve 243
 Material loss 405
 Matrix, traffic/trip distribution, network
 101, 103, 104, 419
 Measurement 326, 332, 333, 362, 461
 Mechanical stabilization 257
 Mechanistic pavement design/model
 295–302, 425
 Mechanized surveys 378–81
 Meeting sight distance 212, 213
 Mesa 167
 Metamorphic rocks 165
 Method-specification 155, 314, 330
 Method statement 336
 Mexico 165
 Microcomputer Accident Analysis Package
 (MAAP) 66
 MicroStation 242
 Mike II 179
 Millennium Development Goals 10, 483
 Miner's law 298, 302
 Minorities 116
 Minor road 114, 115, 443, 446–8, 455, 456;
 appraisal 127–32
 Mission statement 26, 29
 Mix design 258, 261
 Mixing plant 262, 278–82
 Mix-in-place 262
 Mobility 10, 14–15, 36, 127, 470, 471
 Mobilization 333, 335, 347
 Modal share/split 54, 102–6
 Model, road investment *Chapter 21*
 Modernization theory 6
 Modified bitumen 267, 268
 Moisture content 149, 152, 154, 173, 249,
 258, 261, 286, 287, 295, 330
 Monitor 71, 73, 85, 87, 89, 90, 110, 337,
 366, 389, 457, 481, 482, 490, 498
 Monopoly 123, 452, 454
 Montmorillonite 172, 173
 Morocco 171
 Morphology 186
 Moss 242
 Motor cycle/cyclist 72, 422
 Moving observer count 376
 Mozambique 18
 Multi-criteria analysis 88, 111, 128–30
 Multilateral aid/assistance 484, 485, 487
 Multilateral Investment Guarantee Agency
 (MIGA) 487
 Murram 166
 Mutual exclusivity 125, 430
 NAASRA *see* National Association of
 Australian State Road Authorities
 Nairobi 181
 Namibia 448, 489
 National Association of Australian State
 Road Authorities (NAASRA) 205
 Nepal 62
 Net output 61
 Net present value (NPV) 93, 125–7, 418,
 430, 431; cost 126–7
 Network, referencing 375
 Network, road 92, 371
 New Engineering Contract (NEC) 315
 Newly industrialized country (NIC) 8
 New Zealand 24, 175, 445, 453, 454
 Nigeria 264
 Nitric oxide, nitrous oxide 83
 Nkrumah 6
 Node 422
 Noise 83–5, 87, 89, 429; barrier 84
 Non-government organization (NGO) 87,
 484, 494
 Non-linearity, stress/strain 297
 Non-motorized transport 17, 19, 38, 42,
 111, 116, 127, 128, 131, 132, 208, 220,
 417, 422, 430
 Non-pay item 333
 No-overtaking 63, 70
 NovaCad 242
 Nuclear density gauge 150, 154
 Nyerere 7
 Objective, target 27, 29, 65, 69, 95, 106,
 109, 110, 451, 457, 462, 468, 498
 Ochrosol 166
 Official development assistance/finance
 (ODA/ODF) 483–6, 490, 491, 498
 Oil, petroleum 7, 82, 106, 120, 264, 447
 On-the-job training 463
 Operation, project/asset *Chapter 20*, 115,
 373, 390
 Opportunity cost 122
 Organic, soils 83, 156, 166, 167
 Organism 79
 Organization 441, 450–7, 464, 493;
 supervision 325–8
 Organizational policy 26–8
 Organization for Economic Co-operation
 and Development (OECD) 4, 86,
 484, 487
 Origin and destination (O–D) 101, 111
 Outfall 413
 Outlet 195, 196, 202; control 197–9

- Overlay design, pavement 302–4, 401
 Overloading 71
 Overpass 80
Overseas Road Note 31 285–92
 Overtaking 214, 218, 232
 Owner, ownership 110, 316, 443, 448, 452, 455–9, 491
- Pakistan 63, 70, 73, 74
 Papua New Guinea 70–2
 Parastatal 39, 76, 452
 Parking 51
 Participation *see* Community participation
 Particle size distribution 148, 155, 157, 167, 176, 253–6, 276
 Particulates 82, 83
 Parties to contract 316, 317
 Partnering 315
 Parts, spare 120
 Passability 206
 Passenger 51, 111
 Passenger car equivalent (E) 44, 49
 Passenger car space equivalent (PCSE) 45
 Passenger car unit (PCU) 44
 Passenger transport forecast 102–6
 Passing place 216, 218
 Passing sight distance 214
 Patch, patching 398, 399, 405
 Pavement 96, 117, 118, 137, 162, 168, 171, 173, 178, 191, 228, 252, 268, 271, 282, 327, 355, 356, 375, 376, 417, 419; construction 251–3, 282, 283; design *Chapter 15*; evaluation 302; management system 390, 488; materials 297
 Paving machine, paver 282, 283
 Payment for works 326, 332–6, 352, 362, 363
 Pedestrian 24, 36, 41, 42, 50, 59, 60, 72, 76, 86, 127, 132, 205, 206; crossing 33, 50, 64; refuge 50
 Pedicab 41
 Penetration 264–6
 Perception time 212
 Performance: based agreements/contracts 310; bond *see* Security, tender; grading (PG) 266; indicator 31–3, 450
 Permeability 147
 Permissible velocity 199–200
 Phasing of alignment 234–6
 Philippines 7, 58, 175, 487
 Photo tone 140, 141
 Piecework 362, 263
 Pie chart 33
 Pier 203, 204
 Pilot map 239
- Pipe culvert 195, 414
 Plan, planning *Chapter 6*, 23, 52, 69, 89, 90, 352, 362, 498; network 95–8, 419; road safety 65, 75; strategic 32, 52, 93–5, 373; training 468–75
 Plant 79, 84
 Plant-mix 262
 Plasticity, plastic limit (PL), plasticity index (PI) 149, 157, 169, 173, 253, 255–7, 259, 261
 Point bar deposit 176, 177
 Point-rainfall 180
 Poisson's ratio 296
 Police 58, 62–4, 72–4
 Policy *Chapter 2*, 23–8, 75, 90, 92, 93, 95, 419, 442, 452, 457, 462, 493; based lending 90; formulation 33–6
 Polishing, polished stone value (PSV) 156, 164, 165, 268
 Political reform 5, 33
 Politics 26, 117, 442, 443, 445, 446, 448, 452
 Pollution, pollutant 79, 82, 83, 85
 Polymer modification 267
 Pond, ponding 410
 Population 5, 14, 106, 111, 348, 352, 429
 Porosity 155, 175
 Porphyries 163
 Pot-hole 33, 257, 271, 282, 376, 398, 403
 Poverty: index 12, 128; line 11; the poor 15–20, 111, 116
 Poverty reduction strategy/paper (PRS/P) 483, 491, 492
 Pozzolan 259, 261
 Premixed asphalt 273–83
 Preparation: project 495–8; works 373
 Prequalification 319, 353
 Present serviceability index/rating (PSI/PSR) 289, 290, 294
 Prevention, accident 66–7
 Price 122–4, 346, 366, 443, 445, 448; adjustment/index 333, 334; control 8; distortion 123
 Prime 282
 Priority 24, 36, 66, 75, 76, 93, 109, 125, 129, 386–8, 430, 431, 455, 467, 468
 Private finance 449
 Private sector 442, 449, 452, 453, 456, 458, 461, 492
 Privatize, privatization 450, 452, 454, 461
 Procedural specification 314
 Procedure 23, 90, 350, 353, 363, 451, 447, 457
 Proctor 151, 153, 288, 331
 Procurement *Chapter 16*, 36, 390, 442, 455, 457

- Producer surplus 128, 131
 Productivity 347, 349, 351, 362–4, 384–6
 Programme 52, 90, 92, 106, 108–10, 336, 373; analysis 416–18, 432–5; training 471–4
 Progress monitoring, control 336, 337, 498
 Project: analysis, appraisal 416, 417, 431, 432, 497; cycle 115, 494–500; execution methods 309–13
 Proposal, technical, financial 317, 318, 475
 Publicity 63, 64, 73, 74
 Public sector/service 8, 441, 456
 Public transport 51, 92, 103, 104
 Purchasing power parity (PPP) 5
 Pycnometer 154
 Pyroclastic 163, 164

 Quality: assurance 317, 330; based selection (QBS) 318; control 310, 329–32; cost-based selection (QCBS) 318; of life 5, 11
 Quarry 82, 84, 242
 Quartering 147
 Quartz, quartzite 165, 167, 170
 QuickBird 141

 Radius, curve 226–9
 Rainfall 17, 170, 171, 179–83, 186, 287
 Rain gauge, autographic 179
 Rank, ranking 127, 386
 Rate of return, internal (IRR) 70, 92, 125–7, 418
 Rate of spread, binder and chippings 270–2
 Rational method 187, 188, 192
 Reaction time 212
 Record 69, 326, 343, 375, 378
 Recruitment 461
 Recurrence interval 181, 188–90, 196
 Recurrent funds 443
 Red coffee soil 175
 Reduction, accident 70
 Regional factor 293–4
 Register, roads 375
 Registration 353
 Regravelling 405, 408
 Regulation 26, 31, 64, 90, 111, 312
 Reliability 285, 293, 294, 297
 Relocation 116
 Rent, rental 364–6
 Requirements analysis 390
 Resettlement 86
 Resident engineer (RE) 325, 326, 331, 336, 338, 339, 342
 Residual value 116
 Resilient modulus 294
 Resource allocation 442, 447, 457

 Response time 179
 Responsibility 443, 457, 458
 Restructuring 441, 453–6
 Retaining wall 82, 248
 Retention money 335
 Return period 181
 Revenue 313, 365, 366, 443, 446–9
 Reynolds number 203, 204
 Rice 7
 Riding quality *see* Roughness
 Right-of-way 79, 85
 Ring and ball 265
 Rip-rap 82, 202, 203
 Risk, analysis 116, 118, 189, 190, 246, 293, 297, 312
 River 81, 237, 239; plain 176; training 204
 Road: density 14, 15; development 13, 14; fund 14, 34, 36, 445–8, 456; industry 21; maintenance benefits 120–1; pricing 445; speed hump 74, 218; users 14, 29, 34, 38, 72, 89, 93, 96, 124, 206, 284, 372, 417, 421, 428–30, 442, 446, 458
 Roads board 14, 448, 457
 Roadside: areas 407–9; check 72
 Rock *Chapter 9*, 84; crushed 255, 256
 Roller, rolling 272, 273, 283, 330
 Rolling thin film oven test (RFTOT) 266
 Romania 454
 Rostow, Walther 6
 Roterde 166
 Roughness: drainage 184, 186, 187, 197, 198; road 120, 124, 131, 132, 206, 282, 284, 302, 376, 378, 418, 421, 424, 426, 427
 Rumble strip 218
 Runoff 79, 81, 170, 183, 184, 187, 188, 191
 Rural transport infrastructure 16, 110–12, 127
 Rut 396, 403

 Safety: traffic *Chapter 4*, 24, 29, 41, 93, 208, 284, 394, 395, 417; vest 394
 St Kitts 30
 Salt, Saline soils 170, 171
 Sampling, materials 146, 147
 Sand 170, 171, 176; cone 150, 154; equivalent test 155
 Sanding 395
 Sandstone 164
 Santiago 45
 SATCC *see* Southern Africa Transport and Communications Commission
 Satellite image 137, 141, 142, 239
 Saturation flow 43
 Saudi Arabia 170
 Scenario analysis 95, 109, 116
 Schedule, contract 323

- Scheduled treatments 381, 382
Schist 165
School 72, 131
Scoping, environmental 87, 89
Scour 186, 199–204
Screed 282, 283
Screening 86, 87, 127, 128, 235–7, 280
Seal, bitumen 257
Seat belt 72, 76
Section 93, 375, 418, 421
Sector programme support (SPS) 492
Sector-wide approach (SWAP) 490
Security, tender 320–2
Sedimentary rocks 164, 165
Sedimentation 79
Seismic refraction 146
Selection of contractor 318–21
Selection of trainer 475
Self-help 352
Self-munching 173
Sensitivity analysis 116, 189, 416
Service volume 46, 47
Setting of emulsion 267
Settlement 248
Severance 79, 85, 116
Shadow finance/toll 312, 449
Shadow price 123, 350
Shale 164, 257
Shell 296, 301
Shields diagram 203, 204
Shortest route 106
Shortlist 317
Shoulder 210, 219–21, 407, 408
Shoving 397
Shrinkage 262
Side slope 222, 248
Sierra Leone 176
Sieve 148, 154–6, 175, 177, 275
Sight distance 211–15, 242
Sight line 80
Sign, road/traffic 70, 218
Silcrete 165
Silica 163, 166
Silt, silting 192, 410
Singapore 7
Site meeting 326, 342, 343, 389
Slate 165
Slip, landslide 81, 245, 246, 409
Slope 82, 184, 222, 243, 249, 408, 409
Slope-area method 186
Slope stability 81, 137, 245
Slurry seal 400, 401
Soak 154
Social capital 12
Social redistribution 5
Social service 444
Social sustainability 19–20
Socio-economics 24, 85, 86, 92, 102, 109, 116, 128, 346, 372, 421
Softening point 264, 265
Soft loan 490
Software 335, 417
Soil *Chapters 8 and 9*, 81, 82, 84, 118, 179, 184, 187, 236, 241, 245, 247, 251, 375
Soil classification 157–60
Somalia 170
Soundness 156, 164
South Africa 62, 448
Southern Africa Transport and Communication Commission (SATCC) 205–22
South Commission 8
South Korea 7
Spare parts 366, 371
Specific gravity 149, 154, 155, 169, 276, 277
Specification 314, 315, 325, 390, 447
Specificity 451, 452, 454
Speed 39, 46, 51, 74, 84, 206, 215, 230, 231, 240, 416
Speed-flow 38, 43–50
Spoil 245
Spot improvement 102
SPOT satellite 141
Spraying temperature, bitumen 273
Sri Lanka 58
Stability, bitumen 277
Stabilize 168, 171, 175, 257–63
Staff, supervision 325–7
Staff training *Chapter 23*
Stakeholder 35, 36, 89, 353, 442, 457
Standard 27, 29, 38, 67, 69, 71, 73, 76, 88, 93, 95, 96, 107, 114, 124, 188–90, 199, 346, 416, 417, 422
Standard, geometric 205, 207, 208, 224, 225, 239
Stated-preference 121
Statement, policy 29, 31, 35
Station-year 182
Statistic 31, 51, 58
Statistical quality control 331
Stick diagram 71
Stilling basin 202
Stopping sight distance 212, 227, 232
Stop sign 63, 64
Storm 179, 181, 187, 192
Stormwater 178, 191
Strain 284, 295–301
Strategic Highway Research Program 266, 275, 294
Strategy, strategic 23, 27, 71, 90, 92–5, 460, 461; analysis 417–20, 436, 437

- Stratification 164
 Stream 78, 81; flow 186
 Streetcar 51
 Strength coefficient, pavement 292–4,
 303, 426
 Stress 284, 295–301
 Stripping 156, 157
 Structural adjustment 7, 8, 14, 492
 Structural number (SN) 290–4, 303, 426, 427
 Structures 119, 375
 Sub-base 252–4
 Subcontractor 321, 353
 Subgrade 137, 162, 173, 178, 252, 261,
 284, 287; strength 285–90, 294, 297
 Sub-Saharan Africa 17
 Subsidy 8
 Sudan 172
 Sulphate content 156
 Sulphur dioxide 83
 Superelevation 220–2, 225, 226, 229, 243
 Superpave 275
 Supervision: contract *Chapter 17*, 310, 311,
 353; engineer 317–19
 Supervisor 360, 361
 Supplier 36, 450, 452–4, 457
 Surface: dressing 268–73, 355, 399;
 treatment 268
 Surinam 176
 Survey, traffic 52, 53
 Sustainability 5, 14, 18, 36, 91, 498
 Sustainable livelihoods 11, 13
 Sweden 448, 455, 489
 Switzerland 445
- Tactical 27, 92
 Tailwater 196–8
 Taiwan 7
 Tandem axle 287
 Tanzania 18
 Target price 313, 315
 Tariff 443
 Task-rate 362
 Tax 123, 346, 444–9
 Teaching 72
 Technical assistance/co-operation (TA/TC)
 464, 488, 490, 492
 Technology: appropriate *Chapter 18*, 37;
 transfer 64
 Temperature 258, 261, 281, 298
 Tender 319–21, 335
 Terminal condition 285–7, 294
 Terms of reference 317, 475
 Terrain: classification 142–4; topography
 47, 118, 207, 210, 211, 216, 219, 220,
 230, 239, 246, 347, 354
 Testing 73, 329, 330
- Thailand 63, 166
 Theissen method 180
 Theodolite 328
 Theoretical-mechanistic pavement design
 295–302
 Three-arc curve 240
 Time: of concentration 187; cost/savings
 121, 132, 429; extension of 337, 338;
 travel 103, 104, 231; waiting 104
 Tir 171
 Titration test 171
 Tobin Tax 9
 Toll 312, 313, 445–7, 449, 452
 Tool 356–60
 Topography *see* Terrain
 Topsoil 82, 245
 Tourism 85
 Tractor 358–60
 Trade, liberalization 9, 14
 Traffic *Chapter 3*, 100, 168, 178, 206, 215,
 285–7, 302, 375, 376, 419; annual average
 daily (AADT) 42, 46, 48, 50, 53–5;
 assignment 102–6; baseline 53; calming
 70; control, and devices 394, 414; count
 54, 55, 376; design 287; development
 100; diverted/reassigned 42, 98, 99, 120;
 engineering 42, 50, 51; generated/
 redistributed 42, 98–100, 106, 119, 122;
 island 50; mix 205, 206; normal 42, 98,
 99; signal/light 63, 64, 86, 240
 Trailer 358–60
 Trainees 471
 Training *Chapter 23*, 72, 73, 75, 76, 347,
 350, 351–4, 363, 364, 488; aids 477;
 level 465, 466; needs analysis 466–70
 Training production unit (TPU) 463
 Trans European Network 90
 Transit fee 446, 447
 Transition (economic/political) 9, 10, 20–1,
 454, 462; curve 225, 228, 229, 243
 Transparency 442
 Transport: corridor 90; demand 87, 100,
 101, 110, 111; policy model 25; sector
 24, 90, 92; services 16, 17, 37, 362
 Transportation model 106
 Transport (and Road) Research Laboratory
 (TRL/TRRL) 66, 70, 72, 73, 142,
 205–22, 301, 302
 TRANSYT 45
 Treatment, maintenance *Chapter 20*, 381–3
 Trellis-pattern drainage 139, 140
 Triaxial test 154
 Trip generation and distribution 101, 103
 TRL, TRRL *see* Transport (and Road)
 Research Laboratory
 Tropical Flood Model, Generalized 188

- Tropical soils and rocks *Chapter 9*
 Tunnel 83
 Turning circle 206
 Turning lane 50, 240
 Turnkey contract 311
 Turnout 191, 411
 Twinning 465
 Tyre pressure 43, 53, 296
 Tyres 120
- Unbound pavements 253–7
 Uncertainty 116, 416
 Unconfined compressive strength 258
 Underpass 80
 Unemployment, underemployment 346
 United Kingdom 62–4, 70, 72, 74, 86, 256, 449, 453, 455
 United Nations (UN) 10, 86, 87, 483, 487–9
 United Nations Children's Fund (UNICEF) 488
 United Nations Development Programme (UNDP) 4, 488
 United Nations Economic Commission for Africa 73
 United Nations Fund for Population Activities (UNFPA) 489
 United Nations High Commission for Refugees (UNHCR) 488
 United Nations Relief and Works Agency (UNRWA) 489
 United Nations Transition Assistance Group (UNTA) 489
 United States 9, 46, 62, 63, 246, 264, 266, 449
 Unpaved roads 402–6
 User cost/rate 118, 313, 237, 333
 User, road *see* Road, users
 Utility charge 445
 Utilization 121
- Value 88, 371
 Vapour pressure 168
 Variation order 326, 333
 Vegetation 141, 167, 171, 173, 177, 179, 184, 186, 187, 201, 227, 408
 Vehicle: fleet 422; maintenance 120; operating costs 119, 120, 124, 131, 208, 229, 231, 421, 428–30
 Vented drift 194
 Vertical alignment 224, 229–34, 242, 243, 247
 Vibro-compaction 155
 Vietnam 62
- Vignette 445, 446
 Visco-elasticity 297, 299
 Viscosity 265–7
 Visibility 211, 227, 232, 240
 Visual aid 477
 Void 147, 148, 274, 276, 295
 Volcanic 163, 175, 255
 Volume: 30th-hour (30th HV) 46, 48, 215; capacity ratio 47–9, 418; change, soil 251; design hourly (DHV) 46, 215
- Wages 347, 353
 Wall, head, retaining, wing 412
 Wash drilling 146
 Washington DC 487
 Water: absorption 155, 169; crossings 79, 237; ground 79, 85, 137, 170, 171, 178, 183, 247; quality 79; resources 78, 79, 111; surface 78, 178, 247; table 79, 176, 190, 287, 288
 Wealth 11, 14
 Wearing course 252
 Weathering 156, 175, 177, 255
 Weibull formula 182
 Weighbridge, weigh-in-motion 56
 Weight–distance charge 445, 446
 Well-being ranking 11
 Wetland 79, 80
 Wheelbarrow 357–9
 White paper 31, 34
 Whole life cost *see* Life cycle costing
 Width, road, carriageway 120, 207, 215, 218
 Wildlife 79, 80, 84
 Willingness to pay 61, 121
 Wingwall 195, 202
 Without-project case *see* Do-nothing
 Women 18, 19, 36, 111, 116
 Work: methods 345, 356; output 384–6; programme/schedule 241, 242, 326, 336, 386, 389, 458
 Works contracts 310, 311
 World Bank 4, 7, 13, 52, 62, 86, 87, 90, 125, 126, 129, 131, 446, 483, 487, 498
 World Food Programme (WFP) 488
 World Health Organization 489
- Years-of-life-lost (YLL) 57
- Zambia 180
 Zimbabwe 26, 33, 35, 64
 Zones: no-passing 47; traffic 101–6
 Zwarte stofgronden 175