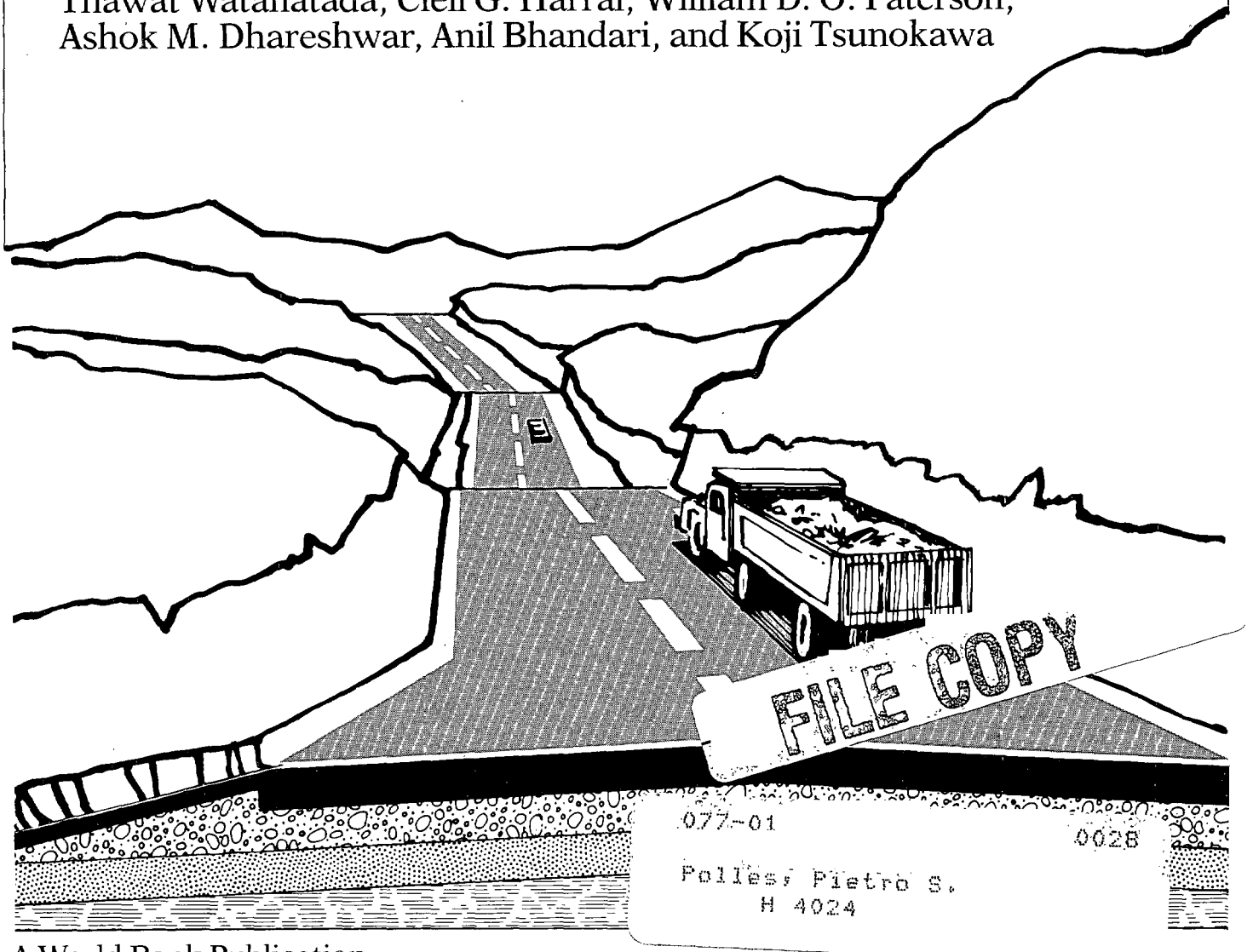
THE HIGHWAY DESIGN AND MAINTENANCE STANDARDS SERIES

9951
VOL. 1

The Highway Design and Maintenance Standards Model

Volume 1
Description of the HDM-III Model

Thawat Watanatada, Clell G. Harral, William D. O. Paterson,
Ashok M. Dhareshwar, Anil Bhandari, and Koji Tsunokawa



A World Bank Publication

12/87

THE HIGHWAY DESIGN AND MAINTENANCE STANDARDS SERIES

The Highway Design and Maintenance Standards Model

Volume 1. Description of the HDM-III Model

Thawat Watanatada
William D. O. Paterson
Anil Bhandari
Clell Harral
Ashok M. Dhareshwar
Koji Tsunokawa

in collaboration with

Wee-Beng Aw
Per E. Fossberg
Edward Holland
Jonathan Rich
Theodora Underhill
Sujiv Vurgese

Published for The World Bank

The Johns Hopkins University Press
Baltimore and London

Manufactured in the United States of America

The Johns Hopkins University Press
Baltimore, Maryland 21211
First printing December 1987

The findings, interpretations, and conclusions expressed in this study are the results of research supported by the World Bank, but they are entirely those of the authors and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent.

Library of Congress Cataloging-in-Publication Data

The Highway design and maintenance standards model.
The Highway design and maintenance standards series)
Bibliography: p.
Contents: v. 1. Description of the HDM-III model—
v. 2. User's manual for the HDM-III model.
1. Roads—Design—Mathematical models. 2. Roads—Maintenance and repair—Mathematical models. 3. Roads—Deterioration—Mathematical models. 4. Motor vehicles—Cost of operation—Mathematical models. I. Thawat Watanatada. II. World Bank. III. Series.
TE175.H54 1987 625.7 87–22733
ISBN 0-8018-3591-7 (v. 1)
ISBN 0-8018-3592-5 (v. 2)

Foreword

An effective road transportation network is an important factor in economic and social development. It is also costly. Road construction and maintenance consume a large proportion of the national budget, while the costs borne by the road-using public for vehicle operation and depreciation are even greater. It is therefore vitally important that policies be pursued which, within financial and other constraints, minimize total transport costs for the individual road links and for the road network as a whole. To do this meaningfully, particularly when dealing with large and diverse road networks, alternatives must be compared and the tradeofffs between them carefully assessed. This in turn requires the ability to quantify and predict performance and cost functions for the desired period of analysis.

Because of the need for such quantitative functions, the World Bank initiated a study in 1969 that later became a large-scale program of collaborative research with leading research institutions and road agencies in several countries. This Highway Design and Maintenance Standards Study (HDM) has focused both on the rigorous empirical quantification of the tradeoffs between the costs of road construction, road maintenance, and vehicle operation and on the development of planning models incorporating total life-cycle cost simulation as a basis for highways decisionmaking.

This volume is one in a series that documents the results of the HDM study. The other volumes are:

Vehicle Operating Costs
Evidence from Developing Countries

Vehicle Speeds and Operating Costs
Models for Road Planning and Management

Road Deterioration and Maintenance Effects
Models for Planning and Management

The Highway Design and Maintenance Standards Model
Volume 2. User's Manual for the HDM-III Model

The Highway Design and Maintenance Standards model resulting from the HDM study is now in its third version, HDM-III. It incorporates the relationships described in the other volumes of this series, as well as a road construction submodel, into interacting sets of costs related to construction, maintenance, and road use. These are added together over time in discounted present values, in which costs are determined by first predicting physical quantities of resource consumption and then multiplying these by unit costs or prices.

HDM-III is designed to make comparative cost estimates and economic evaluations of different construction and maintenance options, including different time-staging strategies, either for a given road project on a specific alignment or for groups of links on an entire network. The user can search for the alternative with the lowest discounted total cost and can call for rates of return, net present values, or first-year benefits. If the HDM is used in conjunction with the Expenditure Budgeting Model, the set of design and maintenance options that would minimize total discounted transport costs or maximize net present value of an entire highway system under year-to-year budget constraints can be determined.

Adequate analysis of the many possible combinations of alternatives is too large a task for manual calculation. Even when analysts have had access to computers of sufficient capacity, they have been hampered by the lack of two essentials: an efficient simulation program embodying an appropriate model and with procedures for using it and an adequate body of empirically established relationships among the relevant variables. The HDM-III model fills both of these needs. It is not only a readily usable program for handling the voluminous computations automatically, it is also a repository of the most extensive and consistent set of empirical data on the subject. The information includes the qualitative structure and

quantitative parameters of relationships among construction standards, maintenance, traffic characteristics, road deterioration, and vehicle operating costs.

This volume describes the HDM-III model and its constituent components and provides a comprehensive discussion of the submodels, their interaction, and the operational parameters involved. A companion volume, *User's Manual for the HDM-III Model,* is essentially for computer mainframe uses, but it also provides the basis for the currently available PC versions.

Clell G. Harral
Principal Transport Economist

Per E. Fossberg
Highways Adviser

Contents

Acknowledgments

This book is the result of a collaborative effort between many individuals and organizations who have contributed directly or indirectly to the various phases of development of the model and to the writing of the book.

First of all, thanks are due to those who established the first model for rigorous analysis to optimize road policies and to minimize total transport cost: F. Moavenzadeh and B. Brademeyer of Massachusetts Institute of Technology. We also thank those who contributed to the viability of the concept and the model through empirical validation: S.W. Abayanayaka, H. Hide, J. Rolt, R. Robinson, and many others at the Transport and Road Research Laboratory in UK; C.G. Swaminathan, L.R. Kadiyali and E. Viswanathan of the Central Road Research Institute in India, as well as the Governments of Kenya and India. Particular thanks are due to the United Nations Development Programme and the Government of Brazil for their sponsorship of the large highway research project which laid the basis for the empirical and theoretical work underlying the major set of relationships in HDM-III. In addition to the numerous contributors to this basic work, special thanks in the context of the HDM-III model are due to T. Lustosa, J. Swait and P.R.S. Resenda Lima of the Brazilian transport planning agency (GEIPOT), and to R. Hudson, B. Butler, R. Harrison, H. Orellana, and L. Moser of Texas Research and Development Foundation; and A. Chesher of Bristol University.

In the writing of the book we have enjoyed the continuous suppot of all the above, as well as many others too numerous to mention. Within the World Bank, we appreciate the continued support of Christopher Willoughby and Louis Pouliquen, respectively the former and present directors of the Transportation Department. Thanks are also due to our numerous Bank colleagues, who over the years have used the various versions of the HDM model, with great patience and understanding for the difficulties encountered in setting up and using a complex model, and who have contributed with suggestions and positive critiques. Finally, we want to express our thanks to S. Lo and R. Archondo-Callao for the excellent research and programming assistance and to E. Danowski, R. Bensky and M. Wu for their great skill and unfailing patience in producing the many drafts that the manuscript went through.

The Highway Design and Maintenance Standards Model

Volume 1. Description of the HDM-III Model

CHAPTER 1

Introduction

1.1 THE PROBLEM

More than 10,000 millions dollars are spent each year on highway construction, maintenance, and administration by governments in the developing countries of Africa, Asia, and Latin America. In the industrialized economies of Europe, North America, and Japan the total is something more than 10 times as great. These numbers are impressive, but it must also be noted that the costs borne by the road-using public for vehicle operation and depreciation are of a much larger magnitude, typically 8- to 10-fold those borne by the government. Moreover, travel time constitutes an additional cost which can assume paramount importance in high-income economies.

In Europe and North America, from which until recently many of the road design and maintenance practices of the whole world have been derived, one sees that factors such as high traffic volumes, high values attached to motorists' travel time savings, and relatively abundant capital resources have dictated high standards of road design and maintenance. With several thousand vehicles per day, even minute savings in vehicle operating costs and travel time can justify very large expenditures on road alignments and pavements. But even in the wealthy countries budgetary pressures on road authorities are now forcing a re-examination of economic priorities.

The situation in the developing countries, which are the focus of World Bank activities, is much different and, of course, far more severe. Relative endowments of capital and labor are much different, traffic volumes are typically much lower, incomes and values attached to travel time savings are far lower, and above all there is an acute shortage of financial resources in general, and foreign exchange in particular. These differences suggest that optimal design and maintenance standards could be quite different in the developing countries. Competing demands for limited resources dictate that low-income countries must search for the most economic design of highways and maintenance programs, taking into account also the much larger costs of vehicle ownership and operation borne by the road users. And even then there will be many economically well-justified projects which cannot be undertaken because of budgetary constraints, so that it becomes imperative to develop a system for assessing the priorities.

But how are we to decide priorities? What is the benefit to society of another dollar spent on maintenance, compared to another spent on new roads or improvements of existing alignments? Is it more economical to spend a bit more money to construct a stronger pavement initially,

thereby permitting the use of larger, more economic vehicles and saving future outlays on road maintenance or, alternatively, should we follow a stage construction strategy, economizing on the initial construction, butrestricting vehicle axleloads and paying more in the way of maintenance and upgrading costs later on, when possible uncertainties about traffic growth will have been resolved? How much, or how little, should we spend to maintain paved roads, and how much to maintain and upgrade earth and gravel roads? And does it matter much if maintenance outlays are postponed during years of financial stringency?

To address these and similar issues the World Bank in 1969 initiated the Highway Design and Maintenance Standards study which ultimately became a major program of collaborative research involving institutions in several countries to develop a new quantitative basis for decision making in the highways sector. In Phase I of the study, completed in 1971, a team at the Massachusetts Institute of Technology, in conjunction with the British Transport and Road Research Laboratory, the French Laboratoire Centrale des Ponts et Chaussées, and the World Bank, developed a conceptual framework and a first prototype model for interrelating the life-cycle costs of highway construction, maintenance and vehicle operation (Moavenzadeh et al., 1971).

While the conceptual framework was felt to be promising, the accompanying survey of previous research revealed an absence of sound empirical evidence from which the fundamental cost relationships could be quantitatively established. Consequently, subsequent phases of the research have concentrated on empirical quantification involving field collection of new primary data on the underlying physical and economic relationships to ensure that the theoretical models conform to the real world as closely as possible. Four such studies have been carried out up to the present time -- in Kenya, the Caribbean, Brazil, and India -- as described below.

As the empirical validation progressed and experience was gained in applying the earlier versions of the model in highway planning in more than 30 countries worldwide, the Highway Design and Maintenance Series (HDM) model has undergone extensive further development and a companion Expenditure Budgeting Model (EBM) has also been developed (Harral et al., 1979; Watanatada and Harral, 1980a; Watanatada and Harral, 1980b). The current third generation version of the model, HDM-III, draws on all of this research and experience and is the focus of this book.

In the remainder of this chapter we provide a simple conceptual overview of the model: its scope, structure, and operation, a summary description of its theoretical underpinnings and empirical validation, and, in the final section, a brief evaluation of the range of validity of the model and its transferability to diverse physical and economic environments. In Chapters 2-6, we provide a more detailed description of the model, its various submodels (traffic, construction, road deterioration and maintenance, and vehicle operating costs) and the interactions among them. Finally, Chapter 7 delineates the economic analysis features of the HDM model, while Chapter 8 discusses the Expenditure Budgeting Model (not an integral part of the HDM itself) and the interface between the two. The comprehensive Volume 2: HDM-III User's Manual guides the user on model

implementation and everyday usage. For users interested in studying the program structure of the model with a view to altering the code to incorporate results of their further research, a Programmer's Guide (Rich and Underhill, 1987) is available. The reader who wishes to examine the theoretical and empirical foundations of the model in greater depth is referred to the other volumes in the Highway Design and Maintenance Standards Series.[1]

It is important to note at this point five important limitations of the present HDM model: first, the vehicle operating costs submodel has not yet been validated for congested traffic conditions. Second, the road deterioration submodel has not been validated for freezing climates, nor third, does it encompass rigid pavements. The reader who must deal with any of these conditions must recognize that the HDM model at its current stage of development, if used at all, must be used with caution to avoid misleading results. (Indeed, in using any such model outside the specific physical and economic context in which it was estimated requires caution and attention to local calibration, a point to which we return in the final section of this chapter.) Fourth, the model does not endogenously address the issue of road accidents nor, fifth, broader environmental impacts such as air or noise pollution, although these costs, where known, can be exogenously incorporated. Thus the model in its present form will be of little interest to the analyst who is primarily concerned with the analysis of rigid pavements, congestion, safety or environmental interventions. Nor is the model intended to be used for final engineering design, rather it is a tool for economic analysis of alternative standards, either at the project or network level.

1.2 THE HIGHWAY DESIGN AND MAINTENANCE MODEL (HDM)

To construct and maintain the road network highway, authorities must choose from a wide range of options, involving the initial standards of pavement and roadway alignment and the frequency and standards of subsequent routine and periodic maintenance, pavement strengthening and geometric improvements. Closely related public policies on vehicle size and weight limits must also be determined. These choices, in turn, have a strong influence on the cost of vehicle operation and thereby the cost of freight and passenger transport. The number of alternative design/policy combinations is very large and decisions taken today will influence operations and costs for years to come.

---

1 Vehicle Operating Costs: Evidence from Developing Countries, Chesher and Harrison (1987).
Vehicle Speeds and Operating Costs: Models for Road Planning and Management, Watanatada et al. (1987).
Deterioration and Maintenance Effects: Models for Planning and Management, Paterson (1987).

Thus, the basic task is to predict total life-cycle costs -- construction, maintenance and road user costs -- as a function of the road design, maintenance standards and other policy options which may be considered.[2] To have a generally applicable tool, one must know the effects of different environments (terrain, climate, traffic, driver behavior, economic conditions) on the different cost relationships. To search over many alternative strategies to determine the most economic, there must be a capability for the rapid calculation and comparison of alternative cost streams which may extend over many years.

The broad concept of the HDM model, as illustrated in Figure 1.1, is quite simple. Three interacting sets of cost relationships are added together over time in discounted present values, where costs are determined by first predicting physical quantities of resource consumption which are then multiplied by unit costs or prices:

$$\text{Construction costs} = f_1\left[\begin{array}{l}\text{terrain, soils, rainfall; geometric design;}\\ \text{pavement design; unit costs}\end{array}\right]$$

$$\text{Maintenance costs} = f_2\left[\begin{array}{l}\text{road deterioration (pavement design, climate,}\\ \text{time, traffic); maintenance standards;}\\ \text{unit costs}\end{array}\right]$$

$$\text{Road user costs} = f_3\left[\begin{array}{l}\text{geometric design; road surface condition;}\\ \text{vehicle speed; vehicle type; unit costs}\end{array}\right]$$

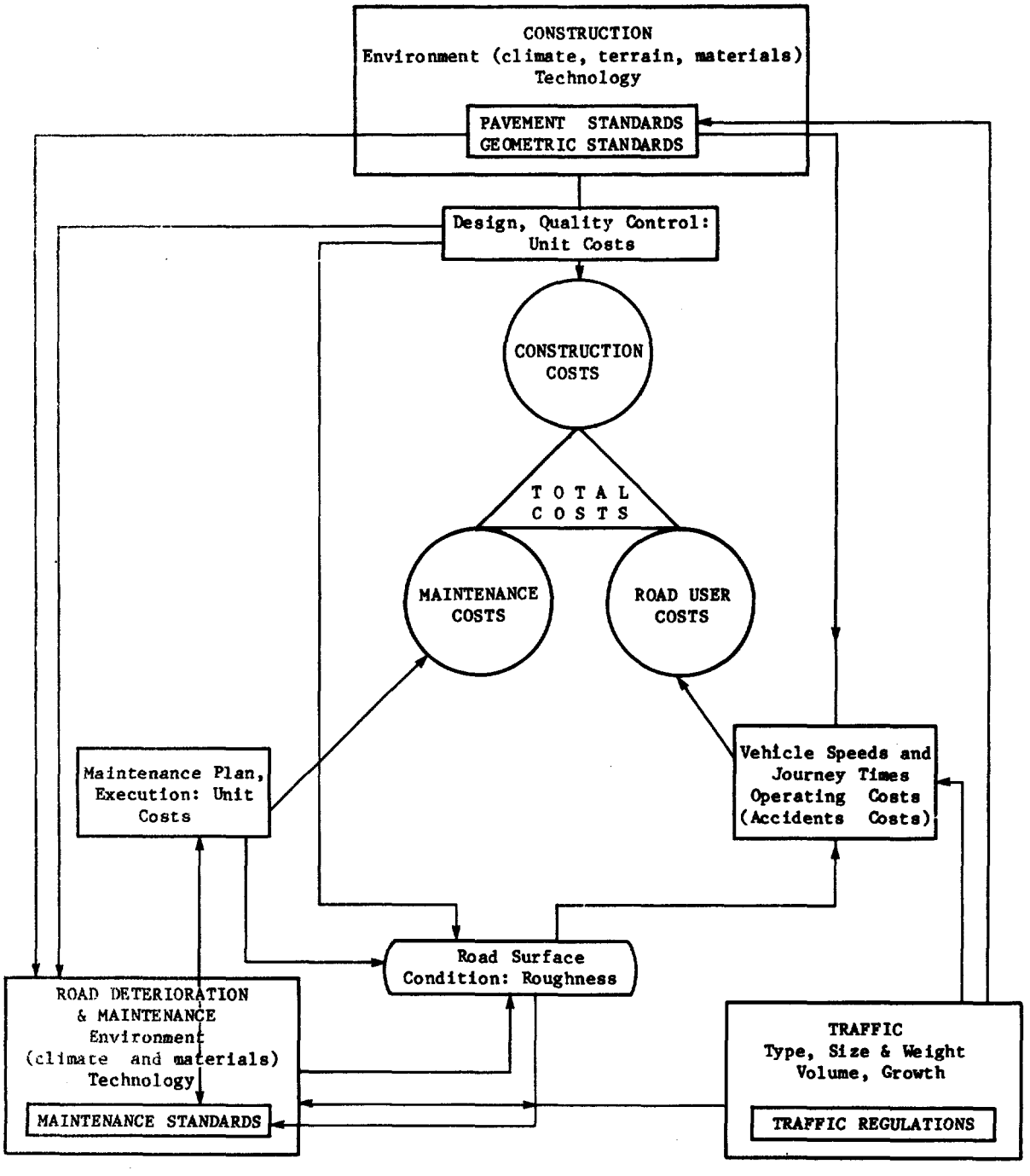
Vehicle speed, which is a major determinant of vehicle operating costs, itself related through a complex set of probabilistic functions to road geometric design, surface condition, vehicle type and driver behavior.

The HDM model is used to make comparative cost estimates and economic evaluations of different policy options, including different time staging strategies, either for a given road project on a specific alignment or for groups of links on an entire network. It can quickly estimate the total costs for large numbers of alternative project designs and policies year by year for up to thirty years, discounting the future costs if desired at different postulated interest rates, so that the user can search for the alternative with the lowest discounted total cost. Or, if he prefers, the user can call for comparisons in terms of rate of return, net present value, or first year benefit. Another capability, using the Expenditure Budgeting Model, is finding the set of design and maintenance options that would minimize total discounted transport costs or maximize

---

2 In certain circumstances an even broader definition of societal costs is necessary, e.g., where the costs of air pollution from road use suffered by non-road-users is significant. Such external costs, may, if known, be entered into the model through the exogenous benefits and costs facility.

Figure 1.1: The HDM Model: Interaction of Costs of Road Construction Maintenance and Use



net present value of an entire highway system under multi-year budget constraints. In addition to comparing alternatives, the model can analyze the sensitivity of the results to changes in assumptions about key variables such as unit cost, traffic growth rates, the discount rate, and the value of passengers' time. A general summary of the scope of the model in terms of input requirements, their limits and the outputs that can be generated is provided in Table 1.1.

As seen from the table, in a single computer run, the model can evaluate up to twenty different road links, each having up to ten sections with different design standards and environmental conditions. Each link can have a different traffic volume. Further, different maintenance standards can be implemented on different sections. At any time, any section can be upgraded (e.g., from earth to gravel or from gravel to paved) and the road can be realigned or widened. Altogether, up to fifty pairs of alternatives can be compared in one run.

In order to make these comparisons, of course, the model must be given detailed specifications of the various alternative sets of construction programs, design standards, and maintenance and other policies to be analyzed, together with unit costs, projected traffic volumes, and environmental conditions. Since there is always the possibility of error in coding these inputs, the model includes an extensive checking program which examines the inputs for formal errors and internal inconsistencies. Warning messages are automatically produced when such errors or inconsistencies are found, or when the program is requested to extrapolate relationships beyond their empirically validated range.

Once the apparent input errors have been corrected, the model estimates speeds and resource consumption of the vehicles as well as road deterioration and resources for maintenance for all the combinations. The resource requirements for road construction for each design option may be endogenously estimated in the model or may be directly specified by the user if he has more specific information or local data. After physical quantities involved in construction, maintenance, and vehicle operation are estimated, user-specified prices and unit costs are applied to determine financial and economic costs. Comparisons in terms of relative benefits, present value, and rate of return calculations then follow. The user has a wide range of options in specifying what results he wants included in the printed report.

Because some of the model relationships have highly complex non-linear forms, simulation, rather than any formal optimization technique, is employed in the HDM model itself, and the "optimization" which takes place in that model is merely the selection of the group of alternatives with the highest discounted net benefits among those specified by the user. There is the possibility that an untried policy combination could exist which would provide superior results to any of those specified for analysis. However, the ease of specifying and analyzing large numbers of alternatives reduces the practical importance of this limitation -- and the alternative approach of simplifying the relationships to a form more amenable to formal optimization could constitute a critical departure from the complex realities of the physical world. When the user comes to

Table 1.1: HDM-III Model Inputs and Outputs

Inputs	Input limits
o Link characteristics (existing road and environmental factors)	20 links
o Construction projects and costs (widening or new construction standards for assigning to links)	50 projects with maximum duration of 5 years for any one project
o Maintenance standards and unit costs (intervention criteria, properties and costs for assigning to links)	30 standards
o Vehicle fleet characteristics and unit costs (common to all link-groups)	9 vehicle types
o Traffic volumes, distribution and growth (sets for assigning to links)	20 traffic sets
o Exogenous costs and benefits (sets for assigning by link)	20 sets
o Link alternatives (assign to links the above construction and maintenance standards, traffic and exogenous C-B sets)	100
o Group alternatives (assign link-alternatives to link-groups)	100 group alternatives involving not more than 20 groups or 100 link-alternatives
o Number of studies, economic comparisons and sensitivity analysis (defines groups to be compared and type of analysis)	Up to 5 studies with maximum group comparisons of 50 with the number of alternatives compared not to exceed 200; 5 discount rates (in addition to zero) per study
o Report requests (uniform per run)	Maximum of 500 reports
o Analysis period (uniform per run)	Up to 30 years with the product of link alternatives and number of analysis years not to exceed 800

Outputs
o Road maintenance summary (by link or group)
o Annual road maintenance costs and physical quantities (by link or group)
o Annual traffic (link only)
o Annual road conditions (link only)
o Annual road user costs and physical resources consumption (link only)
o Financial costs of alternative (link or group)
o Economic and foreign exchange costs of alternative (link and group)
o Comparison of costs of alternatives (link and group)
o Summary of comparison of alternatives by discount rate (link and group)
o Summary of costs and comparisons by discount rate (link alternatives only)

consider the impact of expenditure constraints on the composition of the best feasible group among the alternatives specified, formal optimization techniques, in the form of dynamic programming or heuristic approximations thereto, are available in the Expenditure Budgeting Model.

1.3 EMPIRICAL QUANTIFICATION OF THE BASIC RELATIONSHIPS

Of the three basic sets of relationships -- construction, road deterioration/maintenance and road user costs -- it was evident at the conclusion of the first phase of the research that most of what was needed was already known about estimating construction costs, but far too little was known about the relationships of user costs, road deterioration and maintenance costs to road design and maintenance policies. Consequently, only limited research effort was devoted to filling certain specific needs with respect to construction costs, and, since vehicle operating costs constitute much the largest component of total life-cycle costs (typically 70-90 percent of total costs in low-income countries), the greatest effort was devoted to that subject. Large efforts were also made to establish the road deterioration relationships which are the major determinants of both road maintenance costs and vehicle operating costs. We deal with each of these in turn.

1.3.1 Road Construction Costs

Construction cost estimation is one of the oldest and best established branches of engineering, and earlier versions of the HDM simply provided for construction costs for different alternatives to be exogenously specified by the user -- which is still retained as one option in HDM-III. However, as comprehensive highway sector (or network) level planning assumed increasing importance, the need for an additional, endogenous facility for estimating construction costs was perceived. At the stage of highway sector planning and resource allocation where the range of investment options to be examined is the widest, policy-makers need a method of construction cost prediction that requires minimal information inputs and yet produces cost estimates properly sensitive to a broad spectrum of design standards and terrain characteristics.

After realizing that no such method existed in suitable form, the World Bank and the Massachusetts Institute of Technology initiated in 1980 a small-scale collaborative study to develop a set of relationships for predicting road construction costs that would meet the broad requirements above. As detailed in Aw (1981), road construction data were compiled from 52 road projects located in 28 countries in Asia, Africa, and Central and South America. The regions in which these road projects were constructed cover a broad spectrum of topographic, climatic and soil characteristics -- from flat plains in the Sudan to extremely mountainous areas in Nepal, from the abundance of monsoon rainfall in Pakistan to the dryness of inland Africa, and from areas of good soil materials to lands of poor road-making volcanic ash. The types of construction varied from feeder roads to four-lane freeways, from earth roads to concrete paved roads, and from 30 to 100 km/h design speeds. The first product as reported in Aw (1981, 1982) and Markow and Aw (1983) essentially consisted of a comprehensive data base and a set of preliminary relationships with heavy reliance on

engineering principles in their formulation. These relationships were further refined into a form suitable for general applications, as reported in Tsunokawa (1983). Chapter 3 below provides a summary description of the data base as well as the resulting models and their functioning within the HDM model.

1.3.2 Vehicle Operating Costs and Other User Costs

A decision was made early in the program to concentrate user-costs research on vehicle operating costs relationships rather than the value of time and accident costs. Three factors entered this decision. First, not only are vehicle operating costs the largest cost component in low-income developing countries, but there was already considerable evidence to suggest that they were relatively sensitive to variations in road design and surface condition -- yet none of these relations were then well quantified over a wide range of conditions. Second, as to the value of time, a great deal of complex (but not altogether fruitful) research had already been conducted in high-income countries, where time savings can assume dominant importance, and a review of this literature (Yucel, 1975) indicated that simpler methods could yield approximations to these values that were probably no more unsatisfactory than those yielded by any of the more "sophisticated" methods in use -- and, in any case, in developing countries, which are characterized by low incomes and often extensive un- or under-employment, the issue is much less important. (The HDM model does, of course, endogenously estimate travel times, but the value of time must be exogenously specified by the model user.) Third, with respect to accidents, a complicated set of causalities prevails, with interdependencies among driver behavior and important than road characteristics, and it was felt that within the HDM research program little could be contributed beyond the much larger-scale research efforts already underway worldwide. Moreover, accident costs, although quite large in the aggregate -- as much as 1 percent of GNP in some developing countries (Jacobs and Sayer, 1983) -- are quite small in relation to vehicle operating costs and are in general not as sensitive to variations in road characteristics. Consequently, accident costs are not endogenously modeled in the HDM, but may be specified exogenously where estimates are available.

With respect to vehicle operating costs, major primary research studies were conducted by various institutions in Kenya (1971-75), the Caribbean (1977-82), Brazil (1975-84) and India (1977-83). In this field of research certain key variables can be measured quickly, e.g., vehicle speeds and fuel consumption, while other variables, e.g., vehicle maintenance, can only be observed over longer periods of time. Consequently, each of these studies contained three different components: (1) observations of vehicle speeds under normal usage for a stratified sample of the road network; (2) controlled experiments to establish fuel consumption speed relations as a function of road characteristics; and (3) a user cost survey collecting records over time from large numbers of vehicle operators, for all operating cost components, stratified by road characteristics. Table 1.2 provides summary descriptors indicating the size and range of observations in each of these studies.

Kenya and the Caribbean

The first of these studies was the Kenya study conducted by the British Transport and Road Research Laboratory (TRRL) in collaboration with the Kenya Ministry of Works and the World Bank (Hide et al., 1975). This pioneering study began development of basic measurement methodologies (Abaynayaka, 1976) and was able to establish simple statistical relationships -- mostly in linear forms -- between the various operating cost components (fuel, tires, vehicle maintenance, driver's time, vehicle depreciation and interest) and the principal road characteristics (surface type, roughness, vertical and horizontal alignment) for the somewhat limited range of conditions typical of Kenyan roads. A particularly important discovery was that of the large effect of road roughness on vehicle operating costs for both paved and unpaved roads -- although in the case of paved roads the range of roughness was limited to a value of 4.5 m/km IRI due to the relatively good condition of paved road surfaces which prevailed in Kenya at the time of the study.

Similarly the range of vertical and horizontal geometry was necessarily limited because of the essentially rolling terrain of Kenya. Consequently the relationships for estimating vehicle performance were limited to maximum gradients of 8 percent and maximum horizontal curvatures of 250 degrees per kilometer, and it was not possible to isolate the effects of road geometry on any operating cost component other than fuel. Moreover, since the linear correlation models do not incorporate any of the underlying physical and behavioral processes, and all effects are additive, extrapolation can produce unreasonable results. Indeed, if the Kenya models are extrapolated over higher ranges of vertical and horizontal geometries, as required for the null (or 'baseline') case in incremental benefit-cost analyses of alternative alignments in more severe terrain, they predict negative speeds.

Therefore the smaller-scale Caribbean study was designed and undertaken by the TRRL as a complementary effort to further study the effects of geometry and to extend the range of the relationships for hilly and mountainous terrain and for very rough paved roads. (Hide, 1982; Morosiuk and Abaynayaka, 1982). Because they were conducted by the same team using the same methodologies, observing similar vehicle types (except that buses were not included in the latter study), the Kenya and Caribbean studies provide a good basis for comparison and an opportunity to evaluate how different physical and economic environments affect the basic relationships.

The speed observations and controlled fuel experiments were conducted only on the island of St. Lucia, with rolling to mountainous terrain, where vertical gradients up to 11 percent and horizontal curvature up to nearly 1,100 degrees per kilometer were observed. The user cost survey encompassed also the islands of Barbados, Dominica, and St. Vincent, with varying terrain characteristics, permitting some stratification in the sample. All are small islands, so that the trip length and average annual kilometers utilization by the vehicles are quite low in comparison to other countries.

Table 1.2: Scope of Primary Studies on Vehicle Operating Costs

	Kenya	Caribbean	Brazil	India
A. User survey				
Types of vehicles	5	4	5	3
Total vehicles	289	68	1675	939
Largest truck GVW	26	12	40	28
Companies/operators	NAV	NAV	147	121
Length of observations (yr.)	2	2	4	3
Size of road network monitored (km)	9,300	NAV	36,000	40000
Range of route average				
Roughness (m/km IRI)	3.3-9.0	3.5-11.4	1.8-14.9	5.4-12.9
Average rise + fall(m/km)	14.8-69.4	8-68	10-49	5.8-41.3
Horizontal curvature (deg/km)	1.5-49.7	90-1040	6-294	25.6-675.3
B. Speed observation studies				
Sites	95	28	108	102
Types of vehicles	5	4	6	6
No. of observations	NAV	38,000	76,000	14,000
Range of specific links				
Roughness (m/km IRI)	2.1-22.1	2.0-14.6	1.6-12.2	2.8-16.9
Vertical gradient (%)	0.1-8.6	0-11.1	0-10.8	0-9.1
Horizontal curvature (deg/km)	0-198	0-1099	0-2,866	1-1243
Road width (m)	3.5-7.9	4.3-8.5	5.5-12.9	3.5-7.0
C. Controlled experiments: Fuel				
Test sections	95	82	51	NAV
Types of vehicles	3	3	9	5
Vehicle type	1	1	1 or 2	1
Observations	NAV	1,161-2,296	1,192-5,344	104-411
Range of specific links				
Roughness (m/km IRI)	2.1-22.1	2.0-14.6	2.1-13.3	2.9-11.7
Vertical gradient (%)	1.0-8.6	0-11.1	0-13	0-5
Horizonal curvature (deg/km)	.0-198	0-1099	0-340	NAV

Sources: Kenya: Hide et al. (1975); Caribbean: Hide (1982); Morosiuk and Abaynayaka (1982); Brazil: GEIPOT (1982); Watanatada et al. (1987); India: Central Road Research Institute (1982).

Several important conclusions emerge from the comparisons. First, the major influence of road roughness on operating costs noted in Kenya was confirmed in the Caribbean, where it was further shown that these effects are generally larger on very badly deteriorated paved roads (> 4.5 m/km IRI) than would be indicated by a linear extrapolation of the Kenya relationships. Indeed, the rate of spare parts consumption and tire wear on badly deteriorated bituminous roads in the Caribbean was higher than on gravel roads with the same measure of roughness in Kenya (a fact which may be attributable to differences in the exact profile of the surface irregularity which are not discerned by the bump-integrator instrument used for measurement).

Second, with respect to vehicle speeds, the linear models estimated in St. Lucia predict free speeds under ideal conditions (i.e., level, tangent roads) markedly lower than in Kenya, while the reductions in speed from those already lower levels caused by poorer road geometric characteristics are far less substantial, most coefficients ranging from less than 20 percent to just over 40 percent of the Kenyan coefficients. This may be due to actual behavioral differences--that in severe terrain where several factors exist to suppress speed, drivers are less sensitive to an improvement in one factor alone. This type of physical and behavioral phenomenon cannot be captured by simple linear forms which assume constant sensitivity as reflected in additive coefficients.

These results call into question the applicability of the model forms employed in both the Kenyan and Caribbean studies. While the TRRL team (Morosiuk and Abaynayaka, 1982) proposed a simple linear interpolation of the Kenya and Caribbean speed equations to deal with applications of the models in third countries where terrain conditions lie between the extremes observed in Kenya and the Caribbean, other researchers sought alternative approaches, as discussed below, which incorporate theoretical models of the complex physical and behavioral mechanisms involved in determining vehicle speeds.

Brazil

By far the largest of the four major field studies was the Brazil study, conducted between 1975 and 1984 by a joint team of specialists from Brazil and nine other countries.[3] With the full advantage of the results of the Kenya study, and real resources more than five times as large, the Brazil study employed more advanced theoretical and statistical methodologies and generated a far larger data base covering in most respects a broader range of road characteristics and vehicle types better stratified across the factorial design. Moreover, advances were made in measurement methods for the key variable of road roughness, as discussed in Section 1.3.3 below. Finally, this data base was most intensively exploited with successive rounds of analyses over several years leading to successive reformulation of models, and new experiments were undertaken to

3 Financed primarily by the Government of Brazil and the United Nations Development Program, the study was executed by the Empresa Brasileira de Planejamento de Transportes (GEIPOT) jointly with a team from the World Bank and the Texas Research and Development Foundation.

address issues generated from the earlier analyses, as greater understanding was gained. Substantial gains were made, both in model formulation and statistical estimation (GEIPOT, 1982; Watanatada et al., 1987).

One of the major advances in the Brazil research was the development of new non-linear models for predicting vehicle speeds and fuel consumption based on mechanistic and behavioral concepts, in contrast to the simple linear correlation forms employed in both the Kenya and Caribbean studies. By explicitly incorporating the physical mechanisms and governing behavioral constraints in the models through a probabilistic limiting velocity approach, their performance in predicting vehicle speeds over the wide ranges of alternative road characteristics which need to be examined in any incremental benefit-cost analysis is much improved. By relying on probabilistic formulations, they provide a realistic array of actual outcomes in any specific situation--as modelled by either of two micro methods which mimic detailed speed behavior along the heterogeneous road alignment. The statistical methodologies employed also provide aggregate models incorporating the same theoretical properties which permit the use of less detailed average road characteristics, thus furnishing predictions of average speeds and fuel consumption which are still sensitive to the alternative road standards normally considered at the stage of economic analysis (Watanatada et al., 1987). The latter have been incorporated into the HDM-III model.

It is instructive to compare the speed predictions resulting from the Brazilian, Caribbean and Kenyan models. We draw on the normalized results given by Chesher and Harrison (1987). We adopt the supposition that for free speed under ideal conditions and for average speeds each model is the respective best predictor for the specific country in which it was estimated. With respect to free speeds it is interesting to note that speeds in Kenya and Brazil were generally not too different, generally within ± 5 percent except in the case of buses, where Brazilian buses (often long distance express services) travelled at speeds nearly 12 percent higher. Free speeds in the Caribbean, as already noted, were substantially lower. With respect to the effect of different road characteristics two aspects are particularly noteworthy. First, the non-linear Brazilian equations yield slopes with respect to geometric characteristics which are not too different from those of the Kenyan models over the range of the Kenyan models and slopes which are not too different from the Caribbean models over the range of poorer geometric conditions observed there. This is illustrated in Figure 1.2, which graphs predicted speeds of passenger cars against horizontal curvature for all of the models, but only over that range for which each was estimated; a similar though somewhat more mixed pattern holds for other vehicle types on horizontal curvature and for all vehicles on vertical alignment. Thus the Brazil models tend to support the conclusion from the Kenya study that the first deviations from the ideal condition have very large effects and to support the conclusion of the Caribbean study that further worsening of road geometry, once geometry is already bad, has greatly reduced effects. However, as a country of vast land area, free speeds observed on level tangent roads in Brazil were observed to be significantly higher than those in the Caribbean which comprises only small islands on which trips were

necessarily short. Thus the model estimated in Brazil would overpredict the response of speed to small changes in curvature for better roads in the Caribbean (or India, as shown in Figure 1.2 and discussed below), and some calibration, especially of desired speeds, to local conditions would be necessary.

Second, the estimated effect of road roughness on vehicle speeds, as illustrated in Figure 1.3, is quite similar for all vehicles for the Kenyan, Caribbean and Brazilian models over the lower to middle ranges (less than 5 m/km IRI), but over the higher ranges of roughness the Brazil models estimate much sharper effects, which are similar to the estimates of the Indian models. Thus, once again the non-linear forms, reflecting theoretical properties, of the Brazilian models were deemed most appropriate for generalized model formulation.

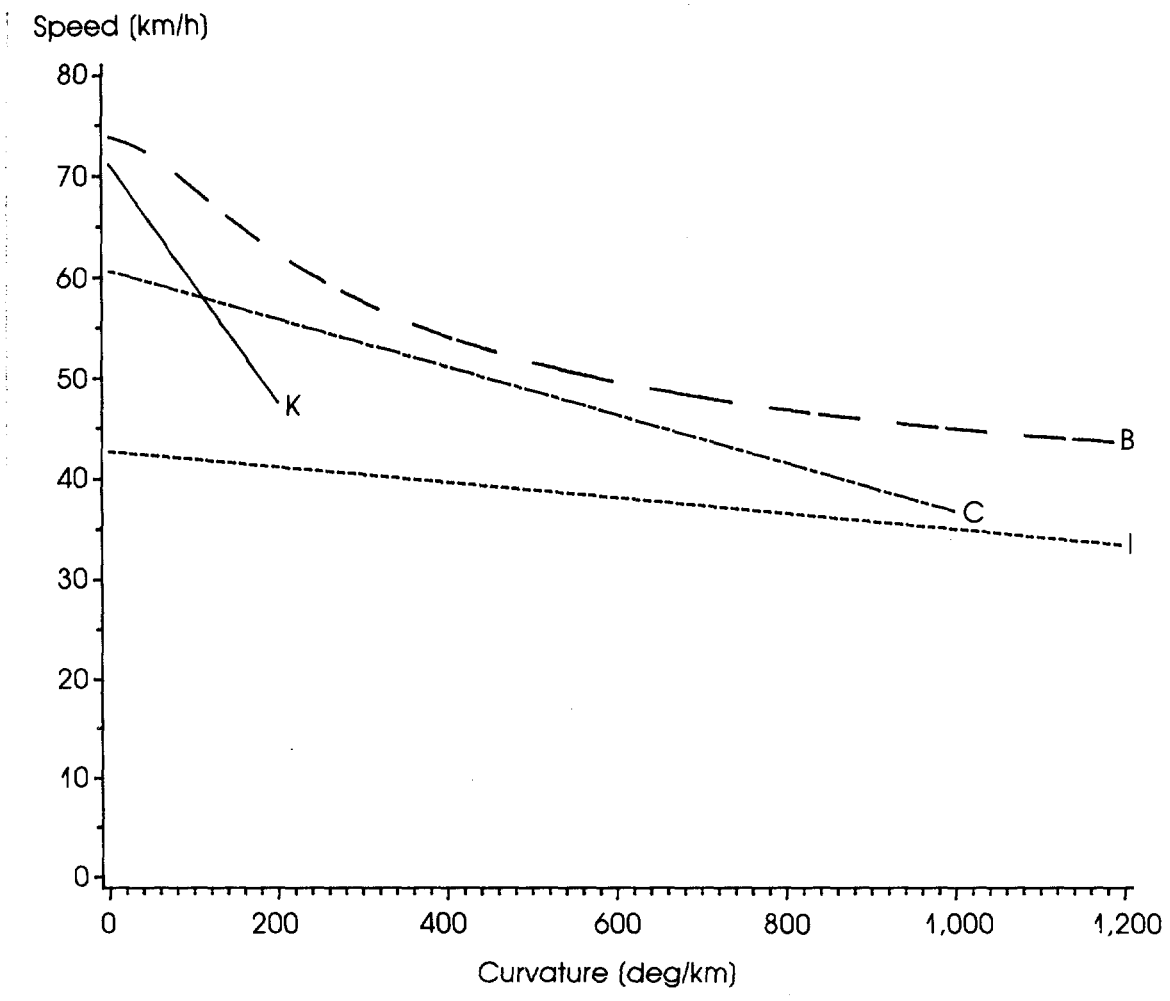
Given the success in estimating these mechanistic-behavioral models for vehicle speeds and fuel consumption, and the strong presumption that other cost components may be physically closely related with vehicle speeds and operating power requirements, the Brazil team was encouraged to extend the same modelling principles to the estimation of tire wear and vehicle maintenance costs. Fortunately, significant research into the physical theories of tire performance had been developed elsewhere by the early 1980s which provided a firm basis for estimation. Despite the fact that the data base in Brazil was ill-structured for such modelling, the resulting mechanistic models for tire wear developed in Brazil yield plausible results. They were deemed to be a significant advance over the earlier linear correlation models, providing a better basis for both present use and future research. However, further research to provide improved models of tire consumption would be desirable, particularly in isolating the respective effects of vertical and horizontal geometry and superelevation, for policy-analysis purposes, and of tire construction type, rubber material, road surface abrasiveness, ambient temperature, and other factors for local adaptation purposes.

Unfortunately, the absence of an accepted body of theory relating the physical wear and tear of vehicles to road characteristics, plus limitations in the data base, prevented any success in estimating mechanistic-behavioral models for that cost component. Consequently, for prediction of vehicle maintenance costs resort had to be made to correlation models. Fortunately, plausible results were obtained with simple model forms for this component, with coefficients with respect to road roughness not too dissimilar to those found in Kenya and the Caribbean (Chesher and Harrison, 1987).

India

There are several unique characteristics of roads and traffic in India which argue against the applicability of models estimated elsewhere. First, much of the Indian network is still single-lane with two directional flows. Second, the traffic mix, although very limited in terms of normal road vehicle types and designs, is extremely heterogenous when account is taken of the multitude of slow-moving vehicle types (agricultural tractors, bullock carts, donkey carts, horse carts, bicycle rickshaws, bicycles,

Figure 1.2: Vehicle speed versus curvature for cars



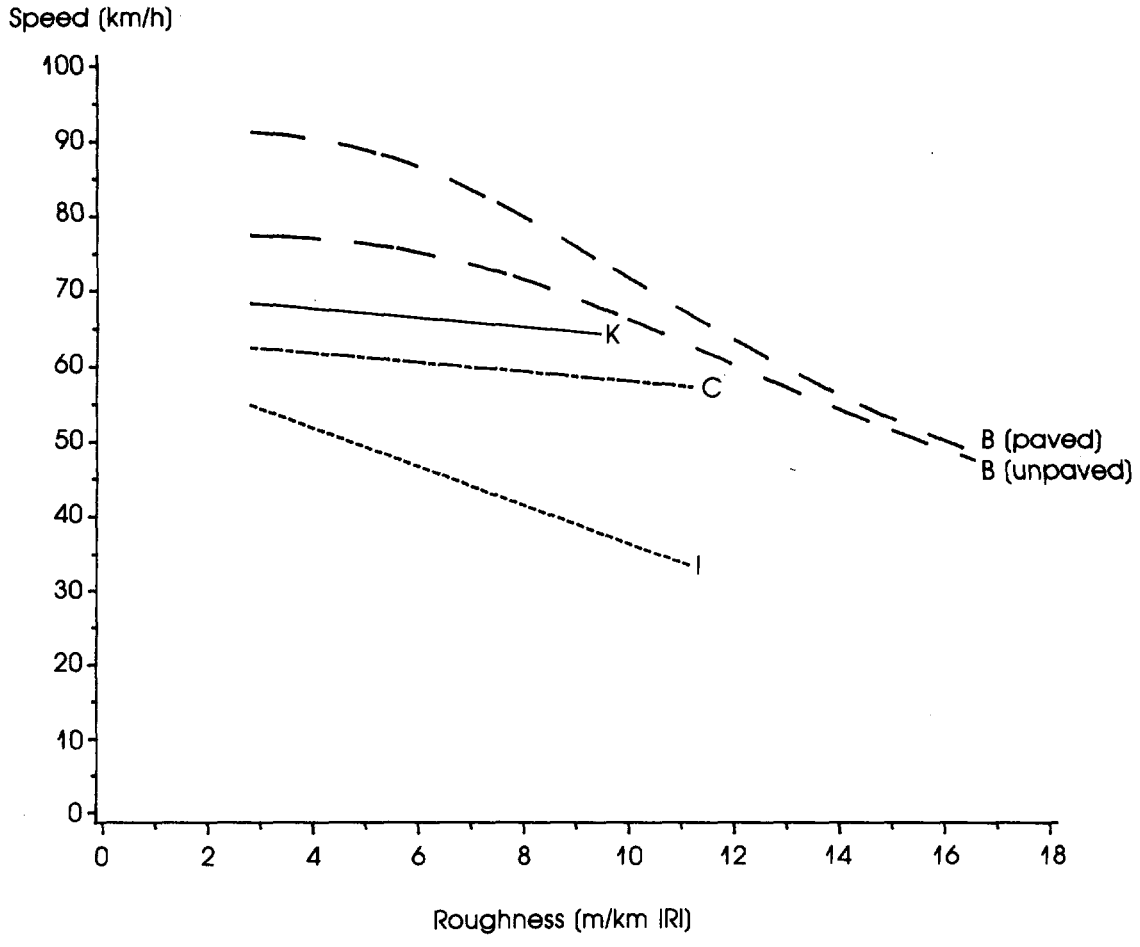
Equations: B = Brazil: Medium cars
K = Kenya: Cars
C = Caribbean: Cars
I = India: Cars

Units: V = Speed (km/h)
RF = Rise plus Fall (m/km)
C = Curvature (°/km)
R = Roughness (mm/km)

Variables not Plotted: RS = 15 m/km
FL = Fall = 15 m/km
R = Roughness = 5,500 mm/km
M = Moisture Content (Kenya only) = 2.6%
RD = Rut Depth (Kenya only) = 18.9 mm
W = Width (India only) = 7 m
ASE = Average Superelevation (Brazil only) = 0.01 (fraction)
ALT = Altitude (Brazil only) = 0
GVW = Gross Vehicle Weight (Brazil only) = 1.4 tonnes

Source: Adapted from Chesher and Harrison, 1987

Figure 1.3: Vehicle speed versus roughness for cars



Equations:
B = Brazil: Medium cars
K = Kenya: Cars
C = Caribbean: Cars
I = India: Cars

Units:
V = Speed (km/h)
RF = Rise plus Fall (m/km)
C = Curvature (°/km)
R = Roughness (mm/km)

Variables not Plotted:
RS = 15 m/km
FL = Fall = 15 m/km
R = Roughness = 50°/km
M = Moisture Content (Kenya only) = 2.6%
RD = Rut Depth (Kenya only) = 18.9 mm
W = Width (India only) = 7 m
ASE = Average Superelevation (Brazil only) = 0.01 (fraction)
ALT = Altitude (Brazil only) = 0
GVW = Gross Vehicle Weight (Brazil only) = 1.4 tonnes

Notes: Results are graphed only for the range covered by the respective field studies.

Source: Adapted from Chesher and Harrison, 1987.

etc.), each with different speed performance characteristics, thus compounding the complexities of traffic flow analysis. Third, the wearing surfaces of paved roads (reflecting indigenous technologies often with insufficient quality controls) tend to be unusually rough, even when relatively new (although normally not as extreme as the badly potholed paved roads observed in the Caribbean study).

In order to address the very different road and traffic conditions in India, the Central Road Research Institute (CRRI-New Delhi) was commissioned in December 1976 by the Government of India and the World Bank to undertake a Road User Costs Study. This project encompassed not only the three basic studies of vehicle speeds, controlled experiments on fuel-speed relations, and a comprehensive user cost survey, but also included pilot studies on simulation modeling of congested traffic flows, road accident costs, and the value of time savings (CRRI, 1982). A rich data base was collected, but, although extensive analyses were performed, it was not possible within the resources and time available to the project to exploit this data base fully. There is considerable instability in the coefficients for given factors across alternative model forms, which results in ambiguities and contradictions. Moreover, further data collection is needed to extend the range of validation of the very promising traffic flow simulation models which have been developed. Consequently the results obtained so far must be viewed as highly tentative, and further research is deemed essential. Indeed, it is hard to avoid the conclusion that the process of understanding and modelling the complexities of vehicle speed-flow and operating cost relationships under Indian conditions has only just begun.

Nonetheless, some important aspects are already quite clearly established. First, road roughness is once again shown to be a major determinant of vehicle operating costs; the importance of establishing and maintaining good riding surfaces is unquestionable. Second, even under the best free-flowing conditions Indian vehicles travel at very low speeds, presumably at least in part due to the very low power-weight ratios for all vehicles. Despite the already very low free speeds of motorized vehicles, the presence of non-motorized vehicles severely reduces the speeds of the motorized traffic, so that as a consequence average motor vehicle speeds on rural highways in India are extremely low, undoubtedly amongst the lowest in the world. To bring road transport service up to higher standards would probably require a combination of measures including traffic separation, vehicle designs and road geometry and surfaces to achieve a new equilibrium; improvements in any one dimension alone may yield little advantage -- indeed, improvements in vehicle performance without improvements in the other dimensions would probably result in a significant worsening of road accidents. Third, building on earlier research in Sweden (Gynnerstedt, 1966; Gynnerstedt et al., 1977) and India (Marwah, 1976), an Indo-Swedish Traffic Flow Simulation Model has been developed which incorporates both the single-lane, two-directional road (as well as two-lane and four-lane roads) and the heterogenous mixture of traffic typical of India (Marwah, 1983; Palaniswamy, 1983; Gynnerstedt, 1984; Palaniswamy et al., 1985). The model has been successfully calibrated for a limited range of traffic volumes and terrain types and offers a promising basis for future development.

1.3.3 Road Deterioration and Maintenance Effects

AASHO Illinois Road Test

In the initial version of the HDM model, the prediction of how road condition deteriorates over time under the action of traffic and weather was based very largely on the results of the AASHO Road Test conducted in Illinois, 1960-61, in a partially freezing climate. That test was by far the largest effort ever undertaken to quantify systematically the complex interactions among road deterioration, traffic (comprising several vehicle types, several axle loadings and axle configurations), and composition of the pavement (in several layers of varying material strengths and varying thicknesses). The primary objective was to determine the relationships between the numbers of axle transits of different loadings and the performance of flexible and rigid pavements for the dual purpose of developing satisfactory pavement design procedures to meet the growing demands of traffic and to aid legislators in setting user taxation and controls for vehicle size and weight. The test was also the first occasion when the many facets of pavement condition and its progressive change over time, which we term "road deterioration," were defined and quantified in a comprehensive statistic, the "pavement serviceability index" (PSI). That index embraced engineers' subjective evaluations of the remaining life of a pavement and its need of maintenance, and was highly correlated to the roughness and, to lesser degrees, to the rut depth and the area of cracking and patching.

However, the applicability of the AASHO test results to roads in developing countries is severely limited by several factors. First, the freezing environment of the test, which had a major influence on deterioration, is distinctly different from the tropical and subtropical climates of most developing countries. Second, the range of pavement types, which was limited primarily to heavy asphalt concrete and rigid surfacings on one weak subgrade, was not representative of the thin surfacings (predominantly surface treatments) and range of material and subgrade types (particularly tropical soils) which are common in developing countries; nor was any gravel or earth road surface included. Third, it is uncertain how applicable the relationships -- based on accelerated, experimentally controlled loading -- are to roads with mixed light and heavy traffic and roads with low traffic volumes. Fourth, in order that different maintenance actions, intervention criteria and standards could be evaluated, it is desirable to predict the trends of roughness, rut depth and cracking separately rather than in the composite serviceability index; this has been encouraged also by more recent developments in the mechanistic theory of pavement behavior and in pavement management. Fifth, the effects of alternative maintenance policies on deterioration were not considered in the AASHO test.

Kenya and Brazil

The road deterioration studies that were conducted in Kenya, 1971 to 1974, and in Brazil, 1977 to 1982, were designed therefore to collect data on the changes of roughness, cracking and rut depth of paved roads in non-freezing climates, over a wide range of pavement strengths and mixed

traffic loadings, and under different maintenance standards. By nature, the rate of change in paved road deterioration over time is both small, because road pavements are usually designed to remain in serviceable condition for 15 to 20 years, and variable, because material properties have high variability. Thus, the four-year study periods were at best the minimum necessary to achieve adequate resolution in the data for the development of predictive models for paved roads, even when the combination of cross-section and time-series methods, as discussed below, is used.

The deterioration of unpaved roads, which constitute a large and important share of the network in developing countries, had received little research attention prior to the Kenya and Brazil studies. The inclusion of wide ranges of both gravel and earth roads in the two studies provided the first basis for an economic evaluation of upgrading to a paved road, and of alternative grading and gravel resurfacing strategies. Since unpaved road deteriorate much faster, results can be obtained more quickly, and both studies extended over periods of approximately two years, encompassing a range of maintenance strategies.

The scope of the two studies is compared in Table 1.3. In the Kenya study, most of the pavements studied were of cement-stabilized base construction and covered a rather narrow range of strengths of 2.7 to 3.7 modified structural number. The volumes and loading of traffic, however, were sufficiently high on eight of the sections so that almost complete deterioration histories were obtained. The study on crushed-stone base pavements was hampered by the loss of seven sections due to drainage difficulties. While the deterioration rates observed agreed reasonably well with current pavement design criteria, the data base was very narrow and the resulting relationships did not extrapolate well beyond the base, particularly for thin pavements.

In the Brazil paved road study, the number of sections and pavement types was more than double that of Kenya and the range of pavement strength covered virtually the whole range currently used in all developing countries, with the exception of very heavily trafficked pavements carrying more than one million equivalent standard axles per lane per year. Excluded from the study were thick bituminous pavements and inverted-design cemented-subbase pavements which are commonly used for very heavily trafficked pavements; there was also no opportunity to observe water bound macadams and bituminous penetration macadams. In the Brazil study the ranges of pavement age, of roughness and of observed roughness change were also double those of Kenya. Axle loadings per vehicle were generally lower, however. The climates of the two study regions are more different than suggested by the rainfall ranges, Brazil's being classed as humid to wet-humid and Kenya's being classed as arid to dry sub-humid. Neither the Kenya nor Brazil studies encompassed either very low or very high rainfall. Horizontal curvature was varied in Kenya but not Brazil. Vertical gradient ranged from 0 to 8 percent in both studies. Pavement width was not varied in Brazil (constant 7.0 m), and varied only slightly in Kenya (6.0 to 7.5 m).

For unpaved roads, the two studies included similar numbers of sections. Both studies included lateritic and rounded quartzitic gravels,

Table 1.3: Scope of primary studies on road deterioration and maintenance in non-freezing climates

	Kenya	Brazil
Paved roads		
Sections	49	116
Granular-base sections	10	74
Cemented-base sections	39	11
Overlaid sections	0	33
Length of sections (m)	1,000	720
Period of observations (year)	4	5
Observations	NAV	500,000
Traffic volume (veh/day)	323-1,618	73-5,700
Equivalent axles (million/lane/year)	0.012-3.6	0.0003-1.7
Equivalent axles per heavy vehicle	0.2-40	0.08-14
Cumulative equivalent axles (million)	0.004-3.3	0.003-18
Annual rainfall	400-2,000	1,200-2,000
Pavement age	0-14	0-24
Modified structural number	2.5-5.1	1.5-7.0
Deflection (Benkelman Beam) (mm)	0.18-1.12	0.20-2.02
Road roughness (m/km IRI)[1]	2.9-6.0	1.8-10.2
Change of roughness (m/km IRI)[1]	0.3-1.7	0-4.9
Unpaved roads		
Sections	46	48
Gravel roads	37	37
Earth roads	9	11
Length of sections (m)	1,000	320-720
Traffic volume (veh/day)	42-403	18-608
Truck volume (veh/day)	12-136	5-477
Annual rainfall (mm/year)	400-2,000	1,200-2,000
Period of observation (years)	2	2.5
Road roughness (m/km IRI)[1]/	4-17	1.5-29

[1] Roughness conversion is given by

$$QI(\text{counts/km}) = 13\ IRI\ (\text{m/km})$$
$$BI\ (\text{mm/km}) = 630\ IRI^{1.12}\ (\text{m/km})$$

(See conducting part of Section 1.3.4 for further details on road roughness measures.)

Sources: Brazil, GEIPOT (1982); Paterson (1987); Kenya, Hodges et al. (1975).

and in addition the Kenya study included volcanic and coral angular gravels; sedimentary gravels were not included in either study, and only nine to eleven sections of earth roads were included in each. The Brazil study included a slightly greater range of traffic volumes, and much greater range of truck volume. The Kenya region was drier but covered a slightly wider range of rainfall than in Brazil, as noted for paved roads.

Research methodology for road deterioration. The road performance studies in Kenya and Brazil had a common objective, namely, to develop models to describe the performance and deterioration of paved and unpaved roads with structural compositions typical of the respective countries, as functions of regional design and construction standards, environmental factors, traffic loading and maintenance policies. Both studies addressed the problem in a similar fashion, although the details of the relationships derived turned out to be quite different.

Excluding controlled load tests (as in the AASHO study) -- which were not attempted in either the Kenya or Brazil studies -- there are two primary methods which can be used to study the performance of existing roads (Hodges, Rolt and Jones, 1975). First, the complete deterioration history of a sample of road test sections can be obtained by monitoring the test sections from the initial construction to their ultimate "failure" (time-series analysis). The main problem encountered in using this method to study the performance of paved roads designed to carry low traffic volumes, or indeed any existing road which has been designed for a long life, is that observations need to be made over a period of many years if complete deterioration histories of the roads are to be obtained. A second method of study is to sample the road population at any instant of time and to include in the sample a representative collection of roads at different stages of their lives (cross-section analysis). The advantage of such cross-section analyses is that results can be obtained much more quickly; the disadvantage of this method is that the data are more scattered because details of the different standards achieved during initial construction and the subsequent deterioration histories of the roads are invariably difficult to obtain. A combination of the two methods was used in both the Kenya and the Brazil studies.

A major principle of both studies was to study road performance under normal operating conditions, rather than through experimental testing. This has several important advantages. First, roads as normally constructed are more representative of the network than experimental sections, which tend to be more closely controlled and, hence, are unlikely to be representative of the actual road network which is to be modelled. Second, it permits both time-series and cross-sectional analysis to be carried out as outlined in the paragraph above, which permits obtaining results in a reasonable time, without resort to accelerated loading (where the time effect on pavement deterioration -- a very important aspect -- tends to be distorted). Third, it is much cheaper to use existing roads with normal traffic than to build separate experimental sections, an important practical concern.

On the other hand, the chosen methodology puts very strict demands on analytical and statistical data treatment -- the study design

has to be well defined and it has to allow for adequate replicates for variations that will occur within each matrix cell. Also, it means a larger number of sections under study, which implies more data collection for establishing road and pavement technical parameters and climatic environment, as well as for monitoring pavement performance and traffic (to determine vehicle flow and equivalent axle flow). And this, in turn, calls for a well designed data handling system, which will permit easy data retrieval and effective analysis.

In both Kenya and Brazil, an initial step was to measure the permanent characteristics of each test section. These included the geometric characteristics such as width, rise, fall and curvature. More importantly, the properties of the pavement layers were determined. These included measurement of strength, layer thickness, particle size distribution, density, moisture content, and plasticity. The basic pavement strength indicator in both studies is the structural number, SN, as conceptually derived during the AASHO study. There were slight differences in the evaluation of SN in the Kenya and the Brazil study -- in the latter, structural coefficients for bound materials were based on compressive strength (for cement-stabilized materials), or on stiffness (e.g., resilient modulus for asphalt concrete and asphaltic bound materials). In both studies, however, layer coefficients for natural soils and gravel, as well as subgrade contribution to SN, are based on the CBR test.

The condition of each test section was regularly monitored during the respective studies. Again, both studies assessed the same type of parameters, namely cracking, potholing effects, rut depth, roughness and deflection. But in this respect, there was a notable divergence of instrumentation and of modelling concepts, particularly in regard to the definition and analysis of cracking and potholes, and in the definition of road roughness and instrumentation for measuring it (as further discussed below in connection with the International Road Roughness Experiment).

This eventually led to considerable differences in the formulation of performance prediction equations. The Kenya equations are characteristically of a continuous function type, each distress function being independent of other distress types, in other words, a parallel development in all distress modes. This makes for relatively straightforward functional forms, but it misses some of the causalities involved in road deterioration. The Brazil study builds on the causality of events, but in so doing, introduces formulational discontinuities, which increase the computational effort, as the reader will see in Chapter 4.

Results on road deterioration and maintenance. Both studies have advanced our ability to predict road deterioration under normal maintenance regimes, and in particular have provided quantitative relationships that give reasonable results when extrapolating over the life of a road. The effects of alternative maintenance policies have been well quantified for unpaved roads, and reasonably qualified for paved roads, except that the longer term effects of repeated maintenance on subsequent deterioration need further research, a point to which we return below.

The large size of the data base and the wide ranges of the major parameters give us confidence that the relative, or marginal, influences of those parameters have been well-represented. For example, on paved roads, the predicted effects of pavement strength and traffic loading compare well with pavement design codes that have been implemented and refined over many years. Dominating the studies, however, is a large degree of scatter in the observed data originating primarily from the inherent variabilities of construction quality, material properties, material behavior and, partly, measurement error. This, in addition to the need to condense the many system variables into a few manageable summary parameters, causes rather wide confidence intervals. The variability, which is typically of the order ± 50 to 100 percent of the predicted change in condition for a given road section, will be of no surprise to the highway engineer. Of greater importance to the planner evaluating a road network is the result that the prediction intervals of the mean, that is the average for the network, are considerably closer than that, on the order of ± 20 percent.

For paved roads, strong relationships have been developed from the Brazil study for original pavements under normal maintenance. Important features of the relationships are their incremental form and inclusion of both traffic and time variables. This permits evaluation of the marginal effects of a vehicle transit, and of time and climate which are important issues in user taxation. The roughness prediction represents a particularly important advance and has been well verified within experimental error on five other data bases (Kenya, three in USA, and southern Africa). From the Brazil study, both roughness and cracking predictions have developed beyond simple correlation models, and incorporate most major mechanistic effects. In the cracking and ravelling relationships, however, construction practices and material properties not easily quantified for network analysis were found to have strong effects, and local linear calibration of these models is recommended. Nevertheless, the predictions of the Brazil relationships for the cracking of cemented base pavements for example compared closely with observed data in Kenya. In Brazil as in Kenya, road gradient was not found to have significant effects on paved road deterioration. The effects of road width were not quantified, although in both Kenya and Brazil the highest levels of distress were sometimes observed in the outer wheelpath.

A useful measure of the relative damaging effects of different axle loads was deduced from the analyses of the Brazil data, which is important because the effects are those under mixed loading and long-term aging, unlike the AASHO Road Test and various recent accelerated controlled load studies. The power applied to axle load in the relative damage function was found to vary with the distress type: for roughness and rutting (i.e., deformation modes) a power value of 4 was found to be valid (in agreement with the AASHO power of approximately 4.2), for cracking initiation and progression lower powers in the order of 2 to 4 were found, and for ravelling (which is distinctly surface wear) a power of near 0 was strongly significant. These effects were generally well-determined for the full data set, except in the cases of cracking and rutting where they were slightly less distinct, and in the cases of a few individual sections. Overall, they serve as a very valuable corroboration of past and present

controlled studies on load-damaging effects, and add new evidence of reduced load-effects on cracking and no load effects on ravelling.

The effects of maintenance on the rate of paved road deterioration, and its initial effect on condition, are as yet only moderately well quantified. In the Brazil study, the major differences in behavior before and after maintenance were in cracking but, apart from initial effects quantified in the models, any significant differences in roughness progression were well quantified through the change in strength parameters. Recent follow-on studies in Kenya, however, have shown very strong reductions in the progression of roughness following multiple reseal applications under apparently negligible changes of strength. Further study is thus required on long term effects over several successive maintenance phases.

For unpaved roads, sound relationships have been developed for the prediction of roughness and material loss, based on material properties (rather than material types as in Kenya), geometry, rainfall and traffic. There is, however, an even higher degree of variability of behavior than for paved roads both across sections, and also across maintenance cycles within a section. Roughness is treated as part of a cyclic process of deterioration under traffic and weather, and maintenance grading, so that the average roughness under a long-term 'steady-state' is predicted. An important innovation in the Brazil relationships is the estimation of minimum and maximum potential roughness levels from material (including particle size), geometric and climatic factors. These levels affect the rate of roughness progression as a function of traffic and also the effectiveness of grading maintenance. Both gradient and horizontal curvature affect unpaved road deterioration. The predictions have been verified on independent data bases from Kenya, Ethiopia, Ghana and southern Africa with acceptable results.

The effects of the limited range of observed rainfall in both the Kenya and Brazil studies were negligible on paved roads. On unpaved roads the Kenya study showed small effects on gravel loss, but the Brazil study showed small effects on both roughness and gravel loss. The relatively small rainfall effect in the paved road relationships may not extend to high rainfall, or low-intensity rainfall climates and requires further study. An important related effect was that the amount of cracking was found to affect rut depth and roughness progression, but the effects may be understated for situations where the pavement layers become saturated.

1.3.4 International Road Roughness Experiment

One of the most important findings of the research in Kenya, sustained by the subsequent studies in the Caribbean, Brazil and India, was the major effect of road profile irregularity, or "roughness," on vehicle operating costs. Road roughness is thus the key variable linking user costs to road surface condition, and the magnitude of these effects is such that they have been found generally to dominate decisions concerning choice of surface type, pavement design and maintenance policies. Because of this discovery the importance of developing a standard reference scale for road roughness and standardized procedures for field measurements became clear. Consequently, the International Road Roughness Experiment (IRRE) was

organized in Brazil in 1982 with the participation of leading research institutes from six countries and the World Bank (Sayers, Gillespie and Queiroz, 1986).

As a result of the IRRE, which encompassed seven major types of instrumentation, roughness has now been defined in an International Roughness Index (IRI). The IRI is a time-stable, transferable, absolute measure of the road profile in a wheeltrack, a dimensionless slope statistic (expressed in units of m/km IRI) which represents the effect of that profile on the axle-body motion of a moving vehicle, idealized in a quarter-car simulation termed $RARS_{80}$ (Sayers, Gillespie and Paterson, 1986). The IRI is similar in concept to, and an improvement on, the QI scale developed in the Brazil study. The roughness data on which all the model relationships in HDM-III are based are the calibrated Maysmeter estimates (denoted QI* in the Brazil study reports) of a reference Quarter-car Index of profile measured by a dynamic profilometer, and this is the measure referred to as QI throughout this volume. The Bump Integrator 'scale' used in the other studies is different, being based on a mechanical device which, although standardized, is inherently subject to small mechanical variations.

The IRRE experiment also provides a sound basis for converting between the different scales used in the various studies, an exercise that is considerably complicated by surface type frequency effects which are introduced by the differences in standard measuring speeds. The Bump Integrator (BI) values tend to vary the most in relation to the IRI in this respect as also do some of the indices from the French APL Profilometer. Typically, road profiles that have high amplitudes in the short wavelength range tend to exaggerate these differences, for example corrugated surfaces and rough earth roads.

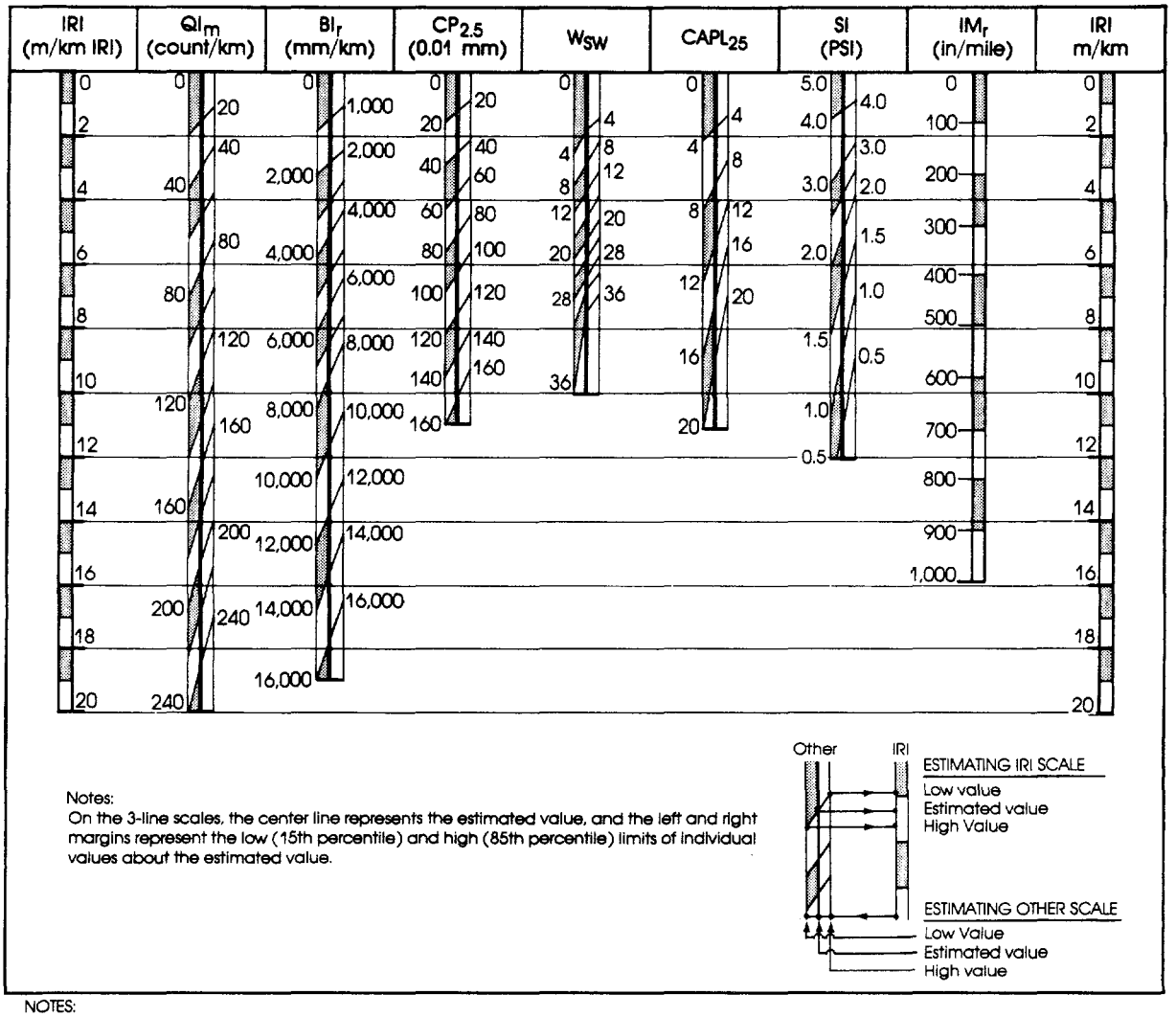
A summary of suitable roughness conversions based on IRRE data given in Figure 1.4 from Paterson (1987); variations in the conversions due to the frequency effects mentioned above are indicated by the range limits shown on the chart. Guidelines for conducting and calibrating roughness measurements to IRI, which is the reference roughness in this volume, are available in Sayers, Gillespie and Paterson (1986). However, here and in the model, roughness in the individual relationships is generally expressed in the scale adopted in the relevant oritinating study, namely QI for the Brazil-UNDP study and BI for the Kenya and India studies. The main coversions are:

$$QI = 13\ RI$$
$$BI = 630\ RI^{1.12} = 55\ QI\ m$$

where RI = roughness, in m/km IRI;
QI = roughness, in counts/km QIm; and
BI = roughness, in man/km BI_r (Bump Integrator Trailer).

The model will accept roughness expressed in various units provided the coefficients of a linear conversion can be supplied, as indicated in Volume 2 HDM-III User's Manual (Chapter 2, Series K, Card K-104).

Figure 1.4: Chart for Approximate Conversions between the International Roughness Index (IRI) and Major Roughness Scales



NOTES:
Conversions estimated on data from the International Road Roughness Experiment, (Sayers, Gillespie and Queiroz, 1986) as follows:

1. IRI — International Roughness Index (Sayers, Gillespie and Paterson, World Bank Technical Paper 46, 1986)
2. QI_m — Quarter-car Index of calibrated Maysmeter, Brazil-UNDP Road Costs Study
   $\hat{IRI} = QI_m/13 \pm 0.37\sqrt{IRI}$ IRI<17
3. BI_r — Bump integrator trailer at 32 km/h, Transport and Road Research Laboratory, UK:
   $\hat{IRI} = 0.0032\, BI_r^{0.89} \pm 0.31\sqrt{IRI}$; IRI<17
4. $CP_{2.5}$ — Coefficient of planarity over 2.5m baselength for APL72 Profilometer, Centre de Recherches Routiers, Belgium:
   $\hat{IRI} = CP_{2.5}/16 \pm 0.27\sqrt{IRI}$; IRI<11
5. W_{sw} — Short Wavelength Energy for APL72 Profilometer, Laboratoire Central des Ponts et Chaussées, France
   $\hat{IRI} = 0.78\, W_{sw}^{0.63} \pm 0.69\, IRI$; IRI<9
6. $CAPL_{25}$ — Coefficient of APL25 Profilometer, Laboratoire Central des Ponts et Chaussées, France:
   $\hat{IRI} = 0.45\, k\, CAPL_{25} \pm 16\%$; IRI<11
   where k = 1 for general use, k = 0.74 for asphalt concrete surfaces, k = 1.11 for surface treatment, earth or gravel
7. SI — Serviceability Index, American Association of State Highway and Transportation Officials:
   $\hat{IRI} = 5.5 \ln(5.0/SI) \pm 25\%$; IRI<12
8. IM_r — Inches/mile equivalent of IRI from Reference Quarter-Car Simulation at 50 mile/hr (see 'HSRI-reference' in Gillespie, Sayers and Segel NCHRP report 228, 1980; and '$RARS_{80}$' in Sayers, Gillespie and Queiroz, World Bank Technical Paper 45, 1986):
   $IRI = IM_r/63.36$

Source: Paterson (1987).

1.4. CONCLUSIONS

The studies in Highway Design and Maintenance Standards Series were undertaken to develop and validate empirically project and sector planning models to permit quantitative analysis of the life-cycle cost tradeoffs in highway construction, maintenance and utilization, and to assess economic priorities. Prior to the initiation of the HDM Series in 1969, and the series of field studies which followed, very little hard scientific evidence was available on the physical and economic inter-relationships among road characteristics, maintenance, and user costs. Consequently, most of the large research effort over the past fifteen years has focussed on empirically quantifying those relationships to ensure that the theoretical models conform as closely as possible to the real world.

Major strides have been made toward this goal, and HDM-III has been developed in a generalized form which can be used with some confidence over a fairly wide range of circumstances to analyze a number of the important investment and maintenance decisions which face road authorities around the world. Because of their more general formulation and stronger statistical quantification, the Brazil models have been chosen as the primary basis for HDM-III. Based on mechanistic principles as far as statistically possible, with explicit incorporation of most of the major causal factors, they are deemed to provide the best basis for interpolation, extrapolation and transference to diverse environments. For road deterioration only the Brazil relationships, as described in Chapter 4, have been incorporated in HDM-III, since they provide the best representation of time and traffic interactions and are also the most suited to studies of marginal effects. For vehicle speeds and operating costs, as discussed in Chapter 5, the Brazil relationships are similarly recommended for most applications, but the Kenya, Caribbean and Indian relationships (presented in Chapter 6) are also included in HDM-III for applications in those particular countries for comparative purposes.

In addressing such a broad and complex set of phenomena over the worldwide diversity of conditions it must be recognized that some factors are better understood, and some relationships are better determined, than others. Life-cycle cost modelling is still in its infancy, and further research will be needed, inter alia: (1) to refine and strengthen the validation of the various prediction models, and to encompass other important phenomena, particularly traffic congestion, and (2) to evaluate the transferability of general model forms and further broaden their empirical validation for diverse physical and economic environments. We conclude this chapter with a brief discussion of each of these issues. The reader who is interested in a more detailed discussion is referred to the other volumes of this series.

1.4.1 Validation of the Models

The most important overall generalization is that the vehicle operating cost models (for free-flowing conditions), and the effects of road conditions thereon, are better determined, and their predictive accuracy is rather greater than is the case for the models for predicting

road deterioration and particularly the effects of maintenance thereon. Various facets of this generalization are touched upon below.

Road user costs

The effects of road characteristics on vehicle operating costs under free-flowing conditions are generally well established for most road conditions and vehicle types in current use. Aside from further analysis of the India data, which merits attention, additional basic research in this area (as distinct from local calibration) is therefore not considered of high priority, although further improvements should be possible with respect to better establishing:

1. The effect of narrow road widths (under 6 m) on speed, fuel consumption, tire and parts wear, and accidents -- an issue of considerable importance in those developing countries where major extensions of the road network are still needed;

2. The effect of small changes in roughness on very smooth paved roads (less than 2.7 m/km IRI) -- an issue of importance primarily in countries where the condition of the network is at very high standards;

3. The effects of superelevation and horizontal alignment on tire wear -- factors whose effects we have not been able to disentangle due to limitations in the data base;

4. The longer-term effects of highway improvements on the utilization and adaptation of the vehicle fleet -- the present models, consistent with traditional practice in highway planning, reflect largely shorter-term impacts on utilization and the model user is left to specify exogenously anticipated longer-term adaptations of the vehicle fleet, a point to which we return below.

With respect to prediction of the effects of road characteristics on vehicle operating costs under congested traffic conditions, very little is yet established other than for speeds and time-related components. congested conditions. A promising start has been made on this problem through collaboration of Swedish, Indian, and Australian researchers using micro-mechanistic behavioral models for traffic flow simulation (as referenced above), and it can be anticipated that further research along these lines will yield satisfactory results in the near future for fuel consumption and possibly tire wear. Satisfactory mechanistic models for vehicle maintenance costs do not appear on the immediate horizon, however, and probably will have to await further developments in the basic theory of vehicle dynamics. In the meantime, since vehicle maintenance costs can be presumed to be sensitive to traffic flow impedance, further research using alternative approaches appears to be warranted.

Road deterioration and maintenance

Performance models to predict the network average behavior of asphalt concrete, bituminous surface dressings, and a range of unpaved road

types have advanced considerably and have now achieved a reasonable degree of accuracy for the subset of worldwide conditions typified by Brazil and Kenya. However, a great deal remains to be done:

1. To improve predictions of the long term effects of alternative maintenance policies for paved roads -- only the immediate effects (e.g., reduction of roughness, cracking, etc.) are reasonably well quantified, and the longer-term effects on retarding subsequent deterioration are quantified for strengthening activities but less well for resurfacings which have had to be adapted from the primary models in HDM-III, with engineering principles and judgment applied to the limited data available. Fortunately, however, it is the immediate effects which are most critical in determining optimal maintenance policies where social discount rates are high (more than 10 percent per annum), and the existing models are expected to yield reasonable answers for such important questions as the choice between resealing and strengthening options, timing of these actions, etc.;

2. To extend the models to a wider range of diverse physical environments, e.g., encompassing very high moisture regimes, high temperature regimes, freezing conditions, etc. This has largely been achieved through the environmental factor values for roughness, but these need verification, and additional parameters may be required in some cases;

3. To extend the models to a wider range of pavement types and construction methods, e.g., to gap-graded materials such as water bound macadams and bituminous penetration macadams (typical of South Asia), to inverted designs, thick asphalt and rigid pavements typically used for very heavily loaded pavements, and, at the opposite extreme, to a wider range of unpaved road types -- although, with respect to the latter, the generalized model form specified in HDM-III, which encompasses material gradation and standard geo-technical parameters, can be expected to transfer reasonably well; and

4. To delineate better the relative damaging effects of different axle configurations and loadings for the purpose of allocating user charges amongst vehicle classes and improving vehicle design. With the advantage of the experience of the HDM studies as well as the AASHO test, further research could usefully now be addressed to this question.

1.4.2 Applicability of the HDM Model in Diverse Physical and Economic Environments

The development of HDM-III has been guided by the objective to develop a general model which could be transferred, with limited local calibration, to diverse countries around the world. While extensive work to quantify the various model relationships empirically under a relatively

wide range of conditions has been done -- in the case of vehicle operating costs through studies in Kenya, the Caribbean, Brazil and India, and in the case of road deterioration relationships through the studies in Kenya, USA and Brazil -- the question remains to what extent can the models be trusted to give accurate answers over the worldwide diversity of physical and economic environments. Essentially the same question can be posed as to the effect of changes over time in the technological and economic circumstances of the countries where the models were originally estimated.

In addressing this issue it is essential to distinguish between the model forms and the model parameters (or coefficients). It is also important to distinguish between the physical and economic environments. Finally, it must be recognized that the ultimate concern is to predict accurately the changes in vehicle operating costs with respect to changes in road characteristics, rather than total operating costs, and cost differentials due to different road conditions are expected to vary much less than total operating costs in response to different economic environments (Chesher and Harrison, 1987).

Where model forms are based on well established theories of physical and behavioral phenomena and incorporate all (or most) of the major determining factors and mechanisms in sufficient detail, such model forms (as distinct from model parameters) should in principle be transferable across diverse environments. The fact that the various studies in Kenya, the Caribbean, Brazil, and India have separately established essentially the same factors playing similar roles -- and generally with the same relative order of magnitude effects -- supports the view that most of the major factors have been incorporated. An important exception is the effect of wider extremes in moisture and temperature, particularly freezing conditions, which were not observed in any of the primary studies upon which the HDM model is based, and which are known from both theoretical principles and other research to have major effects. Thus, the road deterioration model is not expected to be applicable in freezing climates without further theoretical elaboration and empirical validation -- although the vehicle operating cost models could still be utilized with only minor recalibration.

In the HDM model cost components are, with few exceptions, predicted as the product of a predicted physical quantity and an exogenously supplied resource price or unit cost. This separation of physical quantities and prices is obviously essential to permit transferability of the models either across different economies or across time in a given economy. But it is not sufficient to ensure the transferability of the model across different economic environments. That is because certain of the model's coefficients were themselves determined in part by the broader economic circumstances, market conditions and governmental policies prevailing at the time and place of the original studies. For example, the availability and relative cost of capital and labor have a major impact in determining vehicle replacement decisions and the technological mix used for vehicle maintenance -- and it is not surprising that in India vehicles are maintained over longer lives and that vehicle maintenance is much more labor-intensive than that in the other countries. Also governmental policies on vehicle manufacture and

restrictions on imports of new vehicles and spare parts, as well as other specific market conditions, have a major impact on the choice of vehicle type, replacement and maintenance policies.

The manner in which these various issues are dealt with in the context of the HDM-III model is through the user's exogenous determination and specification of a relatively large number of vehicle attributes and model parameters as inputs. Several of the model coefficients are, of course, expected to vary across different environments. Certain data (e.g., resource prices, vehicle technical characteristics, average vehicle life and annual utilization, road subgrade bearing strength, etc.) are easily obtained and are routine input requirements. Other model parameters (e.g., rolling resistance, desired speed and other driver behavior coefficients required for speed predictions) will require some small-scale speed observations and experimentation for a proper local calibration.[4]

The user must also bear in mind that over the longer run vehicle operators have a proclivity to adapt their vehicles and operating rules to changes in road conditions in order to maximize profit and/or minimize costs. They can, within constraints, vary a number of variables to minimize transport cost, e.g., the number, type and make of vehicles, the engine size, age and tires. Of these variables, in keeping with traditional practice in highway planning, only vehicle utilization (or the number of vehicle kilometers driven per year) and, implicitly fleet size,[5] are endogenous within the vehicle operating cost relationships in HDM-III. The treatment of utilization and fleet size as endogenous variables has arisen from their sensitivity to road improvements as well as their considerable influence on vehicle operating cost via the sizable components of depreciation and interest. The other vehicle attributes, once specified, are regarded as constants by the model.

When the user has reason to believe that the changes in road conditions or policies being modelled will lead to other important adaptations of vehicle characteristics, he can incorporate an exogenous estimate of these by specifying a changing composition of the vehicle fleet in the traffic projections. Not to do so would result in an under-estimation of the benefits of an improvement of road conditions or policies, and, conversely, an over-estimation of the losses due to a deterioration, since road users' adaptations will be oriented to improve their position vis-a-vis whatever conditions they face. This is probably not a significant issue in most applications, particularly where the proposed road

---

4 For guidelines on calibration of vehicle speeds and operating cost models see Chapter 13 of Watanatada et al. (1987) and for road deterioration and maintenance models see Chapter 10 of Paterson (1987).

5 Fleet size (number of vehicles employed in a company) is implicitly tied to utilization in the prediction of depreciation and interest costs per kilometer: when the average speed is changed by changes in road characteristics, the number of trips a vehicle can make each year is changed, thereby resulting in a change in the number of vehicles necessary to haul a given volume of transport.

changes are marginal or highly localized, but it could be of some importance where major changes in the road network as a whole or in broad policies, such as axleload limits, were being considered. In the latter cases the model user must give particularly careful attention to specification of the changing characteristics of future traffic types and volumes.

CHAPTER 2

Model Operations and the Traffic Submodel

2.1 BASIC OPERATIONS

The operations of the Highway Design and Maintenance Model take place in three phases; what amounts to a fourth phase is possible by transfering outputs from this model into another model. The first is the data input and diagnostics phase, in which the input data are examined for possible format and numerical errors and internal inconsistencies. Any serious input errors detected in this phase will stop the execution of the remaining phases. The second phase is the simulation of the traffic flows and of the changes in the roads as they go from initial construction through annual cycles of use, deterioration, and maintenance, with possible construction projects to upgrade them. This phase generates information from which, at the user's option, reports may be printed out for specified periods or annually, giving road conditions as well as physical quantities and costs for road construction, road maintenance, and vehicle operation. Benefits to generated traffic and exogenous benefits and costs my be incorporated. The quantities and costs may be broken down, if desired, into components.

The third phase encompasses economic analyses and comparisons of alternative construction and maintenance policies for selected groups of road links. Reports are generated to give differences between the financial, economic, and foreign exchange costs of pairs of alternatives, and compare them in terms of net present value at various discount rates, internal rate of return, and first year benefit.

The other model, with which the HDM may be interfaced, is the Expenditure Budgeting Model (EBM), which selects the optimal combination of projects and maintenance policies under budget constraints.

Each of the three HDM phases and the budgeting model is briefly described in turn below. Following that, the first of the simulation phase submodel -- traffic -- is described in detail. The other submodels are fully described in succeeding chapters.

2.2 DATA INPUT AND DIAGNOSTICS

In order to do all the calculations and comparisons described above, the model must be given detailed data on the roads and vehicles and detailed specifications of the various alternative sets of construction programs, design standards, and maintenance policies to be compared, as well as unit costs, projected traffic flows, and other data. All these data and specifications are punched on cards or coded as card image files,

according to the detailed instructions in Chapter 2 of the Volume 2: HDM-III User's Manual. To help eliminate errors that may have occurred in coding such large amounts of information, the model checks all the input data against built-in criteria to identify departures from prescribed format, inconsistencies, and numerical values outside of expected ranges. Some errors will prevent execution of the simulation, while others are not "fatal". Both types of errors cause messages to be printed out describing the problem encountered. Chapter 3 of the Volume 2: HDM-III User's Manual lists and explains the error messages that may occur.

2.3 SIMULATION PHASE

The sequence of operations of the simulation phase is shown in Figure 2.1. For each year of the analysis period, the submodels shown are applied in succession to each road link with various alternative construction programs and maintenance policies that have been specified for it. (A combination of construction options and maintenance policies that have been designated as a case for analysis on a particular link are referred to as a "link-alternative.")

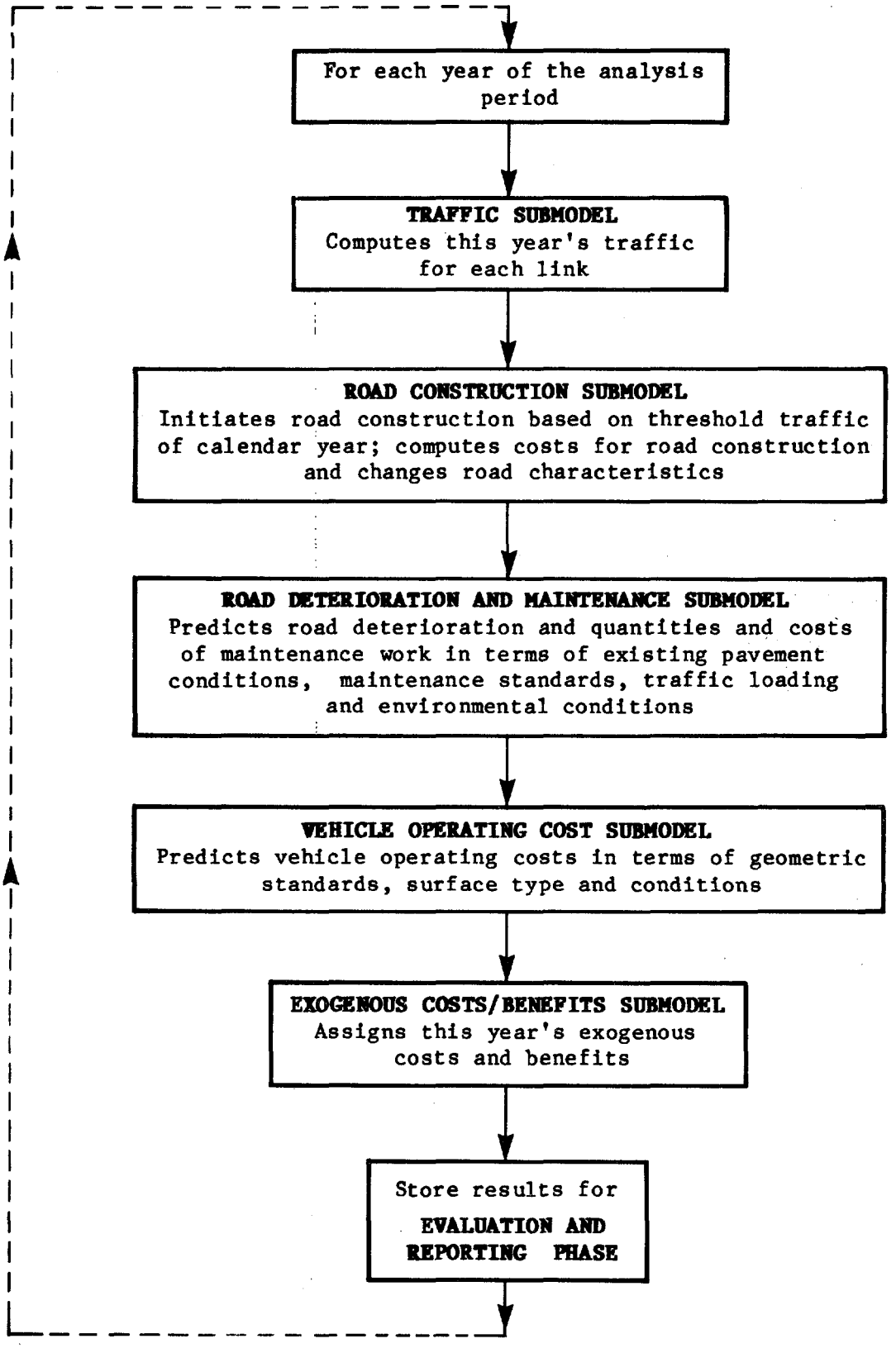
The traffic submodel simply takes data that have been specified by the user in abbreviated form and uses them to calculate the flow of each type of vehicle in each year on each road link. For link-alternatives in which improvements attract additional traffic (relative to the "baseline case") the additional or "generated" traffic must also be specified by the user and added by the model to the "normal" or baseline traffic. The user may specify the time profile of flow of a given vehicle type on a given link either by specifying flow volumes to take effect in particular years and remain constant until changed, or by specifying an initial volume and a rate of growth -- either a fixed increment or a proportion of the current volume each year. Another option, for generated traffic, is to specify it as a fixed ratio to the normal traffic in the same year.

The time profile of a normal (i.e., equal to baseline) traffic flow is specified to begin in a definite calendar year, while the generated traffic is on a relative scale so that it may be initiated by the completion of a construction project, which may be in different years for different alternatives.

The traffic submodel also calculates, for each link-alternative, the number of vehicle axles and the number of equivalent single axles going over the road each year. These values are used in determining the deterioration of the road surface.

For the road construction submodel, the user specifies a baseline schedule of construction projects and as many alternative schedules as he wishes to investigate. A project may be scheduled to begin in a specific year or may be initiated by the volume of traffic reaching a specified level. The duration of each project is also specified by the user and may be from one to five years. Construction projects may include the building of new road links or the widening, realignment, and upgrading of the pavement of existing links.

Figure 2.1: Simulation of a link-alternative



For the baseline case and each alternative, the road construction submodel computes the quantities of work and materials required in each year for the construction projects that have been specified and determines their financial, economic, and foreign exchange costs. Upon completion of a project it changes the physical characteristics of the link involved, and on or before completion it assigns the corresponding generated traffic to it and accounts for any prespecified exogenous costs and benefits.

The total cost per kilometer of a project is made up of various components, and each component cost -- except overhead and "other costs" -- is the product of a physical quantity and a unit cost. One of several optional ways that the model can be used is for the user to specify, on the basis of separate analyses, all of the physical quantities per kilometer and all of the unit costs, as well as overhead and "other." Or, if the multiplications have already been carried out, he can put in the resulting cost per kilometer for each component or even the total for all components, either for an entire link or for separate sections of a link.

Another option makes use of endogenous relations derived from the empirical studies for estimating the quantities involved in certain aspects of construction. These built-in relations, combined with user-supplied unit costs, provide preliminary cost estimates in case engineering has not yet been done. They are particularly useful in analyzing tradeoffs among construction, maintenance, and vehicle operating costs for the investigation of construction standards and maintenance policies at the highway sector level.

The road deterioration and maintenance submodel is the key to analyzing the effects of design and maintenance policy on the condition of roads and hence on vehicle operating costs as a component of the total cost picture. This submodel predicts, for each year, the deterioration of the road surface caused by traffic and climate and the extent to which it is offset by work done under the prescribed maintenance policy. It calculates the quantities involved in the maintenance work, and applies unit costs to determine total maintenance cost for each year. The physical effects of deterioration and maintenance are simulated on the basis of empirical relations derived mainly from the Brazil study (GEIPOT, 1982; Paterson, 1987). The statistical and engineering analyses are fully documented in the latter source.

The submodel accounts for deterioration of paved roads in the form of cracking, ravelling, pothole formation, and rut deepening, all of which affect the progression of roughness, which is the measure of road surface condition used in the vehicle operating cost submodel. Unpaved roads deteriorate by becoming rougher and by losing gravel surfacing material. Road deterioration is a function of the original design, material types, traffic volume and its axle load characteristics, environmental conditions, age of pavement, and the maintenance policy pursued. In addition, as shown by the Brazil study, the loss of material from unpaved roads is affected by horizontal curvature, and erosion due to

rainfall is affected by vertical alignment, which also affects the development of roughness from other causes.

Maintenance options to offset the deterioration of paved roads include patching, preventive treatment, resealing, overlaying, and reconstruction. For unpaved roads the options are grading, spot graveling (patching), and gravel resurfacing. In addition, maintenance includes routine attention to drainage, shoulders, and roadside vegetation.

For paved roads, the deterioration and maintenance submodel includes relations dealing with seven different pavement surface types (some of which result from maintenance procedures) and three base types. The relations apply to tropical and subtropical climates, but have not yet been extended to freezing conditions. Two types of unpaved roads -- gravel and earth -- are differentiated by specifying their generic physical attributes.

The vehicle operating cost submodel computes road users' financial costs, economic costs, and foreign exchange costs for each road section for each year. The quantities of resources consumed and times spent in transit are calculated first and then multiplied by prices or unit costs to obtain operating costs and travel time costs. Vehicle speeds and resources consumed -- fuel, tires, vehicle maintenance, etc. -- are related to the volume and composition of traffic, to the surface type and geometric charactristics of the road section (as initially specified or as altered from time to time by the construction program), and to the current roughness of the surface (as determined from the road deterioration and maintenance submodel).

The analyst chooses among four different sets of relationships which were developed in the separate studies done in Kenya, the Caribbean, India, and Brazil. Each set of relations reflects to some extent the road conditions, the vehicle fleet, and the economic environment of the study region. For each study a somewhat different scheme of vehicle classification was used, corresponding to the range of vehicles commonly used in the different countries. The studies also differed in the variety of independent variables observed and in analytical methodology, as discussed in Chapter 1.

Having specified one of the four sets of relationships for the study, the user must provide data on the rise and fall, horizontal curvature, carriageway width, and surface roughness for each section of each link. Some of the sets of relationships also require specifying surface type, altitude, and other factors. Characteristics of the different types of vehicles in the fleet must also be provided, such as vehicle service life, annual utilization, rated power, gross weight, and so on.

For each year of each alternative, the model first calculates operating speed for each type of vehicle on each road link. This is dependent on the road geometry and roughness and on characteristics of

the vehicle. From the speed, hilliness, roughness, and some other factors, fuel and lubricant consumption and tire wear are determined, as well as parts and labor required for maintenance. For commercial vehicles, crew time on the road is inversely proportional to vehicle speed. For calculating depreciation and interest as proportions of vehicle costs, one is given options of treating vehicle life as constant or varying, and different methods are available for calculating the number of kilometers driven per year and overhead costs. All of these elements, having been calculated in physical or "real" terms, are converted into monetary, economic, and foreign exchange costs by multiplying them by the user-specified unit costs or prices.

For some comparisons of alternatives, it is desirable to account for differences in the time spent by passengers in transit and for differences in the time that cargoes take to reach their destinations. Therefore, the model computes those times and, if the user specifies appropriate unit values, will include these time costs in the analysis.

2.4 ECONOMIC EVALUATION AND REPORTING PHASE

After all the time streams of physical quantities involved in the baseline and each alternative for each road link have been multiplied by the appropriate unit costs or prices, detailed reports and economic comparisons are prepared. Some reports are standard and are produced automatically; others are selected by the user from a larger number of options. Results may be compared for alternative programs of construction and maintenance on individual links. In addition, for comparing possible overall programs, alternatives for different links are usually bundled together into "group alternatives".

The analysis of alternatives in the HDM model can be summarized in the following steps:

1. For each link-alternative, the model separately assembles financial, economic and foreign exchange annual cost streams including the costs of: capital investment, recurrent spending, vehicle operation, passenger and cargo delays, as well as exogenous costs. For illustration see reports type 7 (for link-alternatives) in the computer printout of the sample run. (Volume 2: HDM-III User's Manual, Chapter 5.)

2. The annual cost streams for link-alternatives from step (i) are aggregated for each group-alternative. For illustration see reports type 7 (for group-alternatives) in the sample run printout referred to above.

3. For each pair-wise comparison of link-alternatives the annual benefit and cost streams are computed for one alternative relative to the other in terms of: increases in road capital and recurrent costs; vehicle operating cost and travel time cost savings due to normal traffic;

benefits due to generated traffic; exogenous benefits, and total economic benefits. The savings in foreign exchange are also computed. For illustrations see reports type 8 (for link-alternatives) in the sample run.

4. The cost and benefit streams for step (iii) are summarized for each group-alternative comparison. For illustration, see reports type 8 (for group-alternatives) in the sample run.

5. The model then computes for each pair-wise comparison of link-alternatives: the net present value for five discount rates as specified by the user and automatically for the zero discount rate, the internal rate of return, and the first year benefits. For illustration see reports types 9, 10 and 11 in the sample run.

6. Step 5 is repeated for each pair-wise comparison of group-alternatives. For illustration, see reports types 9 and 10 in the sample run.

For sensitivity studies, steps 3 through 6 above are repeated with certain cost streams multiplied by user-specified factors.

2.5 INTERFACE WITH EXPENDITURE BUDGETING MODEL (EBM)

When using the HDM model to predict the total transport costs and net present values of alternatives relative to a base alternative for individual links, the user can, if desired, maximize the net present value for the highway sector under both capital and recurrent expenditure constraints. This is accomplished by interfacing HDM-III with the Expenditure Budgeting Model (EBM) through a special output file as described in Chapter 8. This output file contains for each alternative the annual net economic benefits relative to the base case as well as the annual capital and recurrent financial road costs.

2.6 THE TRAFFIC SUBMODEL

The simulation phase of the model operation begins, as seen in Figure 2.1, with the operation of the traffic submodel. The traffic submodel employs data entered by the user to derive, for each year of the analysis, the volume of traffic of each vehicle type, the number of vehicle axles, and the number of equivalent single axles on each paved road link under analysis. For unpaved links, only total traffic volumes are calculated.

2.6.1 Traffic Volumes

The data are entered in the form of "traffic sets," defined below, for normal and for generated traffic. Normal traffic on a link is the average daily two-way traffic volume (ADT) of each vehicle group

on the link in each year of the baseline or "without project" case.[1] Generated traffic is the additional volume (of each group in each year) that is induced or diverted from other routes as a result of improved road conditions relative to the base case. Thus in the base case the total traffic is the normal traffic and there is no generated traffic. In alternative cases, the total traffic is the normal traffic plus generated traffic, if there is any.

For any link in a given model run, there is only one base case and therefore only one set of normal traffic, which is the same for all alternatives. Several sets of generated traffic may be available for a link -- different sets, e.g., associated with alternative construction projects. Normal traffic is present at all times; generated traffic usually starts at some time later than the first year -- either in a definite year or at the time of completion of a particular project, which may be different in different alternatives. If, as a result of sequential improvements, more than one generated traffic set is applicable to a link in the same year, the volumes are simply added together.

A traffic set defines a volume of flow for each vehicle group, including its variation over a series of years. A set is made up of one or more "growth periods" in sequence, and the parameters specified for a growth period remain in effect until another growth period begins or until the end of the run. For a normal traffic growth period, an initial calendar year must be specified. If it is the first growth period in the set, this will usually be the first year of the run (although it is permissible and sometimes convenient to designate an earlier year). The starting year for the first growth period in a generated traffic set is Year One, and subsequent growth periods in the same set are timed from that point. Putting generated traffic on a relative time scale makes it convenient to start a generated set at different times under different alternatives, or to make its initiation contingent on other events in the run.

Within each growth period, traffic -- either normal or generated -- may be specified in any of the following three ways:

1. The user may specify fixed volumes for each vehicle group for particular years. Each such amount will remain constant for succeeding years until superseded by a new amount or by the beginning of a new growth period.

2. After initial traffic volumes have been established, the user may specify an annual increment for each vehicle group. The fixed amounts will be added each year until new values for the increments are called for or a new growth period begins.

---

[1] When used in the road deterioration model, the total two-way traffic is divided, in the model, by the effective number of lanes, and half of the total is assumed to be going each way.

3. After initial volumes have been established, an annual percentage growth rate may be specified for each vehicle group. Traffic volumes will be incremented by these percentages each year until new rates are put into effect or a new growth period begins.

For generated traffic only, a fourth option is available. Generated traffic in each group may be specified as a percentage of the group's normal traffic. The specified percentages will be applied until superseded.

2.6.2 Axle loadings

For simulating the effects of traffic on the roads, the model makes use of two different measures of axle load -- vehicle axles and equivalent standard axle loads. The latter combines the damaging effects of the full spectrum of axle loading in a common damage-related unit. The measures are computed in the traffic submodel and passed along for use in the road deterioration and maintenance submodel (see Chapter 4).

The term "vehicle axles" is defined as the total number of axles of all vehicles. The number of vehicle axles traversing a given link in a given year is computed as the volume of traffic in each group multiplied by the number of axles per vehicle of the type involved.

For each vehicle group, the number of "equivalent standard axle loads" traversing a given link in a given year is computed as the product of the annual traffic volume of the group and the group's "equivalent standard axle load factor," specified by the user or computed from axle load information. These numbers are summed over all groups to obtain the total number of equivalent standard axles traversing a link in a given year.

The equivalent standard axle load factor, AF, is defined as the number of applications of a standard 80 kN dual-wheel single axle load which would cause the same amount of damage to a road as one application of the axle load being considered.

The equivalent standard axle load factors ($AF4_k$) are computed as as follows, for the different load-spreading (and damage-reducing) effects of grouped axles, single wheels, etc. which are incorporated by varying the standard load, $SAXL_j$, used in determining the loading ratio:

$$AF4_k = \sum_{i=1}^{I_k} \frac{P_{ki}}{100} \sum_{j=1}^{J_{ki}} \left[\frac{AXL_{kij}}{SAXL_j} \right]^{4.0}$$

$$AF2_k = \sum_{i=1}^{I_k} \frac{P_{ki}}{100} \sum_{j=1}^{J_{ki}} \left[\frac{AXL_{kij}}{SAXL_j} \right]^{2.0}$$

$AF4_k$ = the equivalent standard axle load factor based on the equivalency exponent of LE=4.0 for vehicle group k, in ESA per vehicle;

$AF2_k$ = the equivalent standard axle load factor based on the equivalency exponent of LE=2.0 for vehicle group k, in ESA per vehicle;

LE = the axle load equivalency exponent;

P_{ki} = the percentage of vehicles in subgroup i of the vehicles in group k;

I_k = the number of subgroups in vehicle group k;

J_{ki} = the number of single axles per vehicle in subgroup i of vehicle group k (a tandem axle is treated as two separate single axles);

AXL_{kij} = the average load on axle j of load range i in vehicle group k (tons); and

$SAXL_j$ = standard single axle load of axle group type, j:
= 6.60 ton for single-wheel single axle,
= 8.16 ton for dual-wheel single axle,
= 15.10/2 = 7.55 ton for dual-wheel tandem axle,
= 22.90/3 = 7.63 ton for dual-wheel triple axle.
(Note: 1 ton = 1,000 kg ≃ 9.8 kN)

The i-th subgroup of a vehicle type k comprises all those axles of group type j (single or tandem, etc.) carrying loads AXL_{kij} in the i-th range. The standard axle load $SAXL_j$ is determined solely on the basis of the axle group type j.

In the classical formulation derived from the AASHO Road Test, the pavement structural number (unmodified) and the initial and final values of the serviceability index are used in computing the equivalent standard axle load factor, which differs for single and tandem axles (AASHTO, 1974). For the purposes of simplification and because the economic analysis should not be restricted to pre-determined initial and terminal levels of serviceability, the computational method specified in here includes the average power value of 4 on the axle loading in the relative damage function. Given the much greater inaccuracies usually associated with traffic volume forecasting, the high variability of pavement behavior, and the variety of approaches possible for deriving a value from performance data, such an average value is highly appropriate and the most representative of the pavement damaging effect which leads, ultimately to roughness progression.

Analysis of the Brazil-UNDP study data (Paterson, 1987) in fact provided strong validation of this value, based moreover on an evaluation of pavement performance under mixed traffic on pavements with a wide age-range, and under in-service conditions. The analysis showed evidence that the exponent, LE, has a lower value of approximately 2 in some cases for cracking initiation and such a tendency is supported from theoretical mechanistic considerations (Paterson, 1987; Rauhut, et al., 1984). Provision has been made in the input forms therefore for entry of an equivalent standard axle load factor, $AF2_k$, computed with LE=2 as defined above. In this HDM-III version, however, the coefficients in the

performance prediction equations have been standardized to values of LE=4 and LE=0 (in the traffic flow variables YE4 and YAX respectively), because the values of LE were not always consistent across pavement types and distress modes (see Paterson, 1987). The exponent LE=0 attributes equal damaging effect to every vehicle axle (including light-vehicles) independent of axle load.

The number of vehicle axles, i.e. for LE=0, is defined as equal to the number of axles of the vehicle if it is classified as heavy (3,500 kg gross weight or more), and two otherwise.

If the user chooses to enter vehicle axle loads through cards D-203, the endogenous computation assumes $SAXL_j$ to be a constant, fixed value of 8.2 ton. This assumption regards every axle as a single dual-wheel axle and is valid only when the pavement depth is thin relative to the load magnitude (as a rule of thumb, when axle spacing exceeds half the pavement depth or axle loads exceed 3 times the modified structural number). The default value for light vehicle axle loads is provided such that:

$$AF4_k = 0.1^4$$
$$AF2_k = 0.1^2$$

The user should note that the factor $AF4_k$ is an average applying across <u>all</u> vehicles of type k, loaded and unloaded, in <u>both</u> directions on the given link.

CHAPTER 3

Road Construction Submodel

The main purposes of the road construction submodel are: to compute and allocate construction costs by component (financial, economic, and foreign exchange) on a year-by-year basis over the duration of the construction period; to modify the physical characteristics of the link as the construction is completed; and to activate generated traffic (see Chapter 2) and generated exogenous costs and benefits (see Chapter 7) for the link, if any, upon the construction completion.

Construction may be scheduled for any year in the analysis period and may consist of new construction either on new or old alignment and widening. Construction may also be scheduled sequentially, as in staging; that is, a link may be constructed to a low standard initially, followed at a future date by widening, realignment, or some other form of upgrading.

Major road construction quantities, namely, earthwork, drainage, site preparation and bridges, can, if desired, be endogenously predicted as a function of terrain severity and geometric standards. The relationships used have been developed from data of 52 road construction projects from 28 countries in Africa, Latin America and Asia (Aw, 1981). While these relationships provide preliminary cost estimates where local data are not available, they are particularly useful in analyzing tradeoffs among road construction, maintenance, and vehicle operating costs at the highway sector level. This is because the construction quantities predicted are sensitive to geometric standards, especially for severe terrain (see Section 3.2 for graphical illustrations).

However, the model user can often gain more precise estimates of construction costs where engineering preparation of projects has advanced beyond the preliminary stage. In this case, the user simply specifies construction costs exogenously, as described below.

3.1 BASIC COMPUTATIONAL PROCEDURE

Construction may have a duration of one to five years, and may be initiated in any year during the analysis period by specifying either a starting calendar year or a threshold traffic level. In either way the construction is assumed to be completed at the end of the effective completion year specified by the user. The effective completion year, which is less than or equal to the last year of the construction period, is employed to handle cases in which some portions of the road have already been completed and open to traffic before the completion of the entire construction.

For instance, construction scheduled to start in 1984 with effective completion in year 2 will be initiated at the beginning of 1984, and effectively completed at the end of 1985. It will be available to go into service with new physical characteristics at the beginning of 1986.

The road construction submodel may be divided into the following stages:

1. During the construction. For each year of the construction period, the cost incurred during the year is computed in financial, economic, and foreign exchange terms. A detailed description of cost computation is provided in the next section. Salvage values are also computed by cost component. They are entered as percentages of total financial, economic and foreign exchange costs, and are recorded as benefits in the last analysis year.

2. In the opening year. For the year following the effective completion year, the physical characteristics of the affected link are changed to those of the new road. A detailed description of the requirements for road characteristics data is given in Chapter 2 of the Volume 2: HDM-III User's Manual

In addition, if the user specifies generated traffic and exogenous costs and benefits associated with the road construction, the generated traffic set and the exogenous costs and benefits set are activated and assigned to the link starting in the year following effective completion.

As schematically illustrated in Figure 3.1, road construction costs may in general be broken down into eight components: right-of-way, site preparation, earthwork, pavement, drainage, bridges, other costs, and overhead. Except for other costs and overhead these components may further disaggregate into physical quantities and unit costs. For example, site preparation costs can be obtained as a product of the site clearance area per unit length of road and the cost per unit area. Similarly, earthwork costs can be decomposed into earthwork volume per unit road length and the cost per unit volume.

Taking this structure into account, HDM-III provides a flexible input routine which allows the user to specify the construction costs in different ways (see Table 3.1). Specifically, the user can specify the costs as either total costs (level 1 or 2) or costs of each of the components (level 3). The total costs can be entered either as the costs for the entire link (level 1) or separately for the affected sections in the link (level 2). The user has a further option to enter some or all of the component costs (except other costs and overhead) by the physical quantities and the unit costs (level 4). As mentioned earlier, the quantities of road construction (except the right-of-way area, the pavement volume, and in the case of widening, the floor area of bridges) can either be exogenously specified by the user or estimated endogenously by the model. In the latter case, using the relationships described in Section

Figure 3.1: Construction costs and availability of models for predicting construction quantities

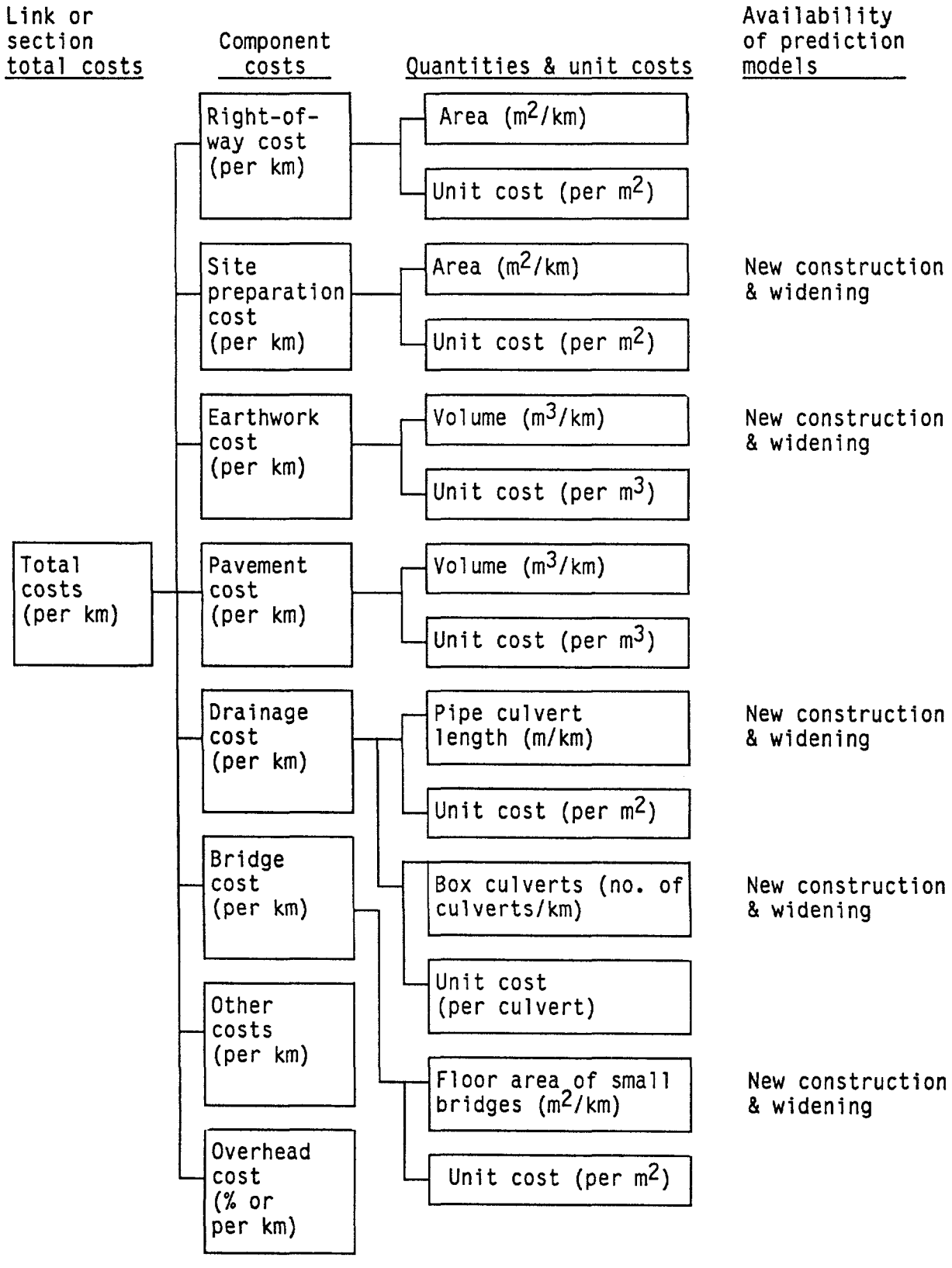


Table 3.1: Structure of Cost Data for Road Construction Submodel

Level	Details
Level 1: Link total cost	The user provides a lump sum total cost for the entire link. The use of this level precludes the use of more detailed levels (2, 3 and 4) for all sections affected.
Level 2: Section total cost	The user provides a lump sum total cost for an affected section on the link. The use of this level precludes the use of levels 3 and 4 for the link, but still permits these levels to be applied to the other affected sections.
Level 3: Component cost ("cost per km")	The user provides the cost per km of a construction cost component for an affected section. The use of this level precludes the use of level 4 (unit cost and quantity) for this cost component and section, but still permits "unit cost and quantity" to be provided for the other components where applicable.
Level 4: Component cost ("unit cost and quantity")	The user provides the unit cost (per unit quantity) and quantity (per km) for a component, for an affected section. Among the eight cost components, the following are applicable: Right-of-way — Pavement Site preparation — Drainage Earthwork — Bridges Of the above components, site preparation, earthwork, drainage and bridgeshave built-in formulas in HDM-III for endogenously predicting the physical quantities.

3.2, they are computed as a function of the geometric characteristics of the road section and the terrain.

"Other costs" are all those which are not otherwise included. The latter may include major structures such as large bridges, large culverts, side drains and so on. "Other costs" are specified on a per kilometer basis.

Overhead costs are often represented as a fixed percentage of the sum of the other component costs. Thus the user can specify overhead costs either as a lump sum cost per km or as a fixed percentage of the sum of the other component costs.

Costs of construction, whether total costs, component costs or unit costs, are input by the user in financial, economic, and foreign exchange terms (though financial and foreign exchange costs are optional). They are distributed over the construction period according to the annual percentages specified by the user. The salvage value -- i.e., value remaining to be realized after the end of the analysis period -- is specified by the user as a fixed percentage of the total construction costs.

3.2 PREDICTING ROAD CONSTRUCTION QUANTITIES

The empirical relationships used to compute the quantities of road construction, viz., the area of site preparation, volume of earthwork, length of pipe culverts, number of box culverts, and floor area of bridges are described in Appendix 3A. The following paragraphs provide a summary of these relationships. The recommended range of the input variables in Table 3.2 should be observed as departure from the range implies an extrapolation and could produce unreasonable results.

3.2.1 Site Preparation

Site preparation involves creating site access and removing trees, stumps, bushes, and obstacles. Depending on whether it is for new construction or widening, the area of site preparation can be estimated according to the relationships described below.

New construction

In the case of new construction on new alignment, the area of site clearing and grubbing per unit length of road is given by:

$$ACG = 1770 \exp(0.0278\ GRF) + 1610 \exp(-0.0114\ GRF)\ RW$$

where ACG = the average area of site clearing and grubbing per unit length of road, in m^2/km;
GRF = the ground rise plus fall, in m/km (see definition below); and
RW = the roadway width, in meters, defined as:
RW = W + 2 WS;
W = the width of the carriageway, in meters; and
WS = the width of one shoulder, in meters.

The ground rise plus fall is defined as the sum of the absolute values of total vertical rise and total vertical fall of the originial ground, in meters, along the road alignment over the road section in either direction divided by the total section length, in km. Figure 3.2 illustrates how this value is computed for a given road section. The predicted area of site clearance is plotted against the roadway width in Figure 3.3.

Table 3.2: Recommended range of variables for road construction quantity prediction[1]

Variable	Units	Recommended range
Roadway width,[1] RW	m	5-25
Road rise plus fall,[2] RF	m/km	0-75
Ground rise plus fall, GRF	m/km	0-100
Rise plus fall differential,[2] G	m/km	0-50

[1] Not used in the drainage relationships.
[2] Used only in the earthwork volume relationships.
Source: Authors' recommendation based on data from Aw (1981;1982).

Widening

In the case of widening, ACG may be estimated by the following incremental formula:

$$ACG = 1610 \exp(-0.0114\ GRF)\ [RW_{(after)} - RW_{(before)}]$$

where the subscripts "after" and "before" represent the values before and after widening, respectively.

3.2.2. Earthwork

The earthwork of a road construction project is that portion of materials required to convert the original ground feature from its natural condition or configuration to the prescribed sections and grades. It includes the quantities of soil, gravel, rock, and unsuitable materials resulting from the excavation, borrow, embankment construction, spoil, and grading operations. Among the major cost components, earthwork is the one which is the most sensitive to geometric standards and terrain conditions.

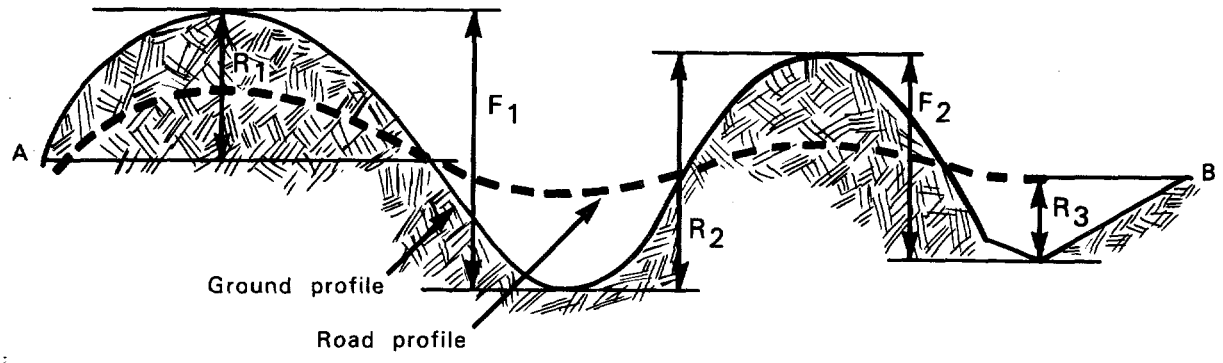
New construction (new alignment and existing alignment)

The following relationship is used to estimate the earthwork volume in the model:

$$EWV = 1000\ (RW + 0.731\ H)\ H$$

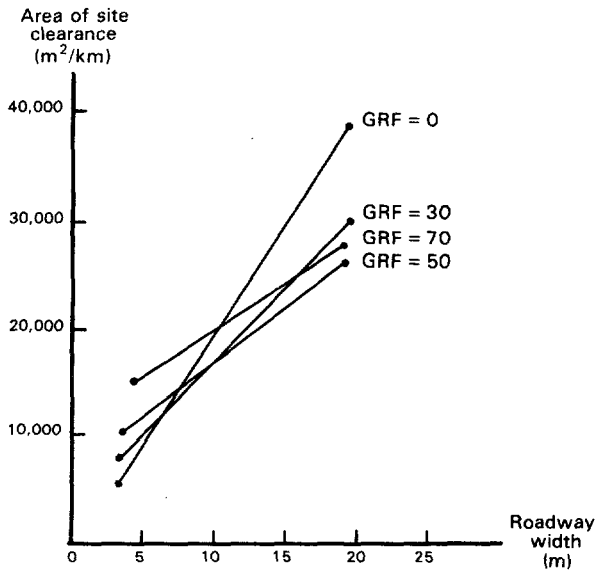
where EWV = the volume of earthwork per unit length of road, in m^3/km (includes cut, fill, borrow and waste materials);
H = the effective height of earthwork, in meters, given by

Figure 3.2: Illustration of ground rise + fall



$$\text{Ground rise + fall} = \frac{R_1 + R_2 + R_3 + F_1 + F_2}{L_{ab}}$$

Figure 3.3: Predicted area of site clearance versus roadway width



Source: Based on analysis in Tsunokawa (1983). See also Aw (1981; 1982) and Markow and Aw (1983).

$H = 1.41 + 0.129\ G + 0.0139\ GRF$; and
G = the rise plus fall differential, in m/km, given by
$G = GRF - RF$

where RF = the road rise plus fall, in m/km.

The road rise plus fall is defined in the same way as the ground rise plus fall except that the vertical profile of the road is used instead (as elaborated in detail in Chapter 5). The sensitivity of the predicted earthwork volume to the roadway width and the road rise plus fall for various values of the ground rise plus fall (representing terrain severity) is shown in Figures 3.4 and 3.5.

Widening

For widening, the earthwork volume is given by

$$EWV = 1000\ H\ [RW_{(after)} - RW_{(before)}]$$

where the variables and subscripts are as defined above.

If GRF falls within the range $0 < GRF < 10$, i.e., flat terrain, the user has the option of inputting the embankment height (in meters) directly as the value for the effective earthwork height, H, in the above expression.

3.2.3 Drainage

Pipe and box culverts are major components of drainage costs. Empirical relationships for predicting the physical quantities of culverts are available for regular pipe culverts (those having diameters between 0.1 and 1.5 m) and box culverts (those having span lengths under 1.0 m). Since the costs of larger pipe and box culverts and side drains, if any, are not estimated by these relationships, they should be included in "other costs."

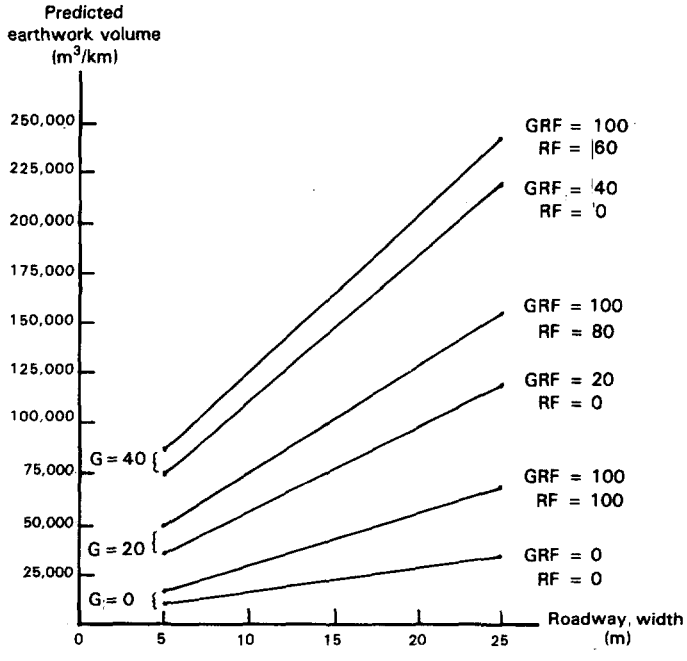
New construction (new and old alignment)

Pipe culverts. The following relationship is used to estimate the aggregate length of regular pipe culverts per unit length of road:

$$DRL = \begin{cases} 1.97\ ALPC & \text{if } 0 \leq GRF < 10 \quad \text{(flat)} \\ 1.74\ ALPC & \text{if } 10 \leq GRF < 40 \quad \text{(rolling)} \\ 2.02\ ALPC & \text{if } 40 \leq GRF \quad \text{(mountainous)} \end{cases}$$

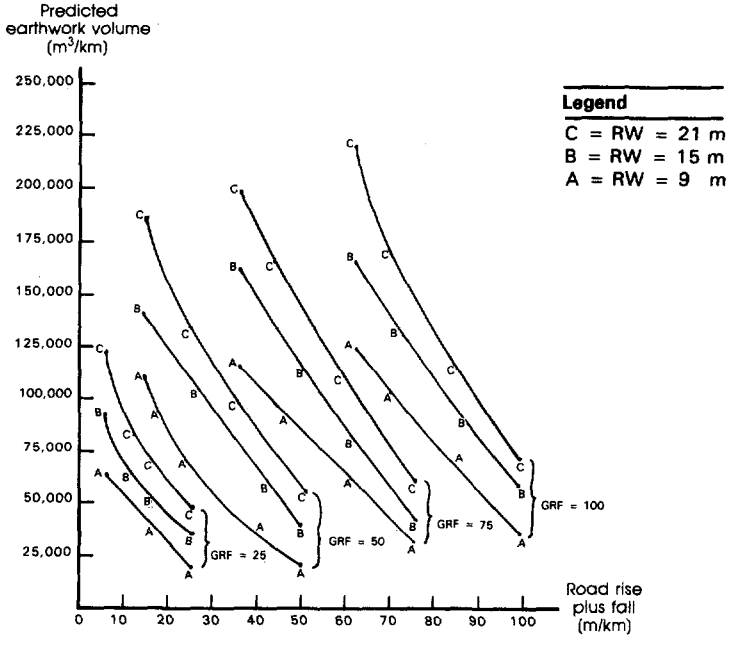
where DRL = the aggregate length of regular pipe culverts per unit length of road, in m/km; and

Figure 3.4: Predicted earthwork volume versus roadway width



Source: Based on analysis in Tsunokawa (1983). See also Aw (1981; 1982) and Markow and Aw (1983).

Figure 3.5: Predicted earthwork volume versus road rise + fall



Source: Based on analysis in Tsunokawa (1983). See also Aw (1981; 1982) and Markow and Aw (1983).

ALPC = the average length of regular pipe culverts, in meters, given by:

$$ALPC = 2.57 \exp(-0.00313\ GRF)\ RW^{0.895}$$

The predicted average length of regular pipe culverts is plotted against the roadway width in Figure 3.6.

Box culverts. The average number of regular box culverts per unit length of road is estimated as follows:

$$ANBC = \begin{cases} 0.27 & \text{if } 0 \leq GRF < 10 \\ 0.15 & \text{if } 10 \leq GRF < 40 \\ 0.62 & \text{if } 40 \leq GRF \end{cases}$$

where ANBC = the average number of regular box culverts per unit length of road, in culverts per km.

Widening

For widening, the pipe culvert length and the number of box culverts are estimated according to the following relationships:

Pipe culverts.

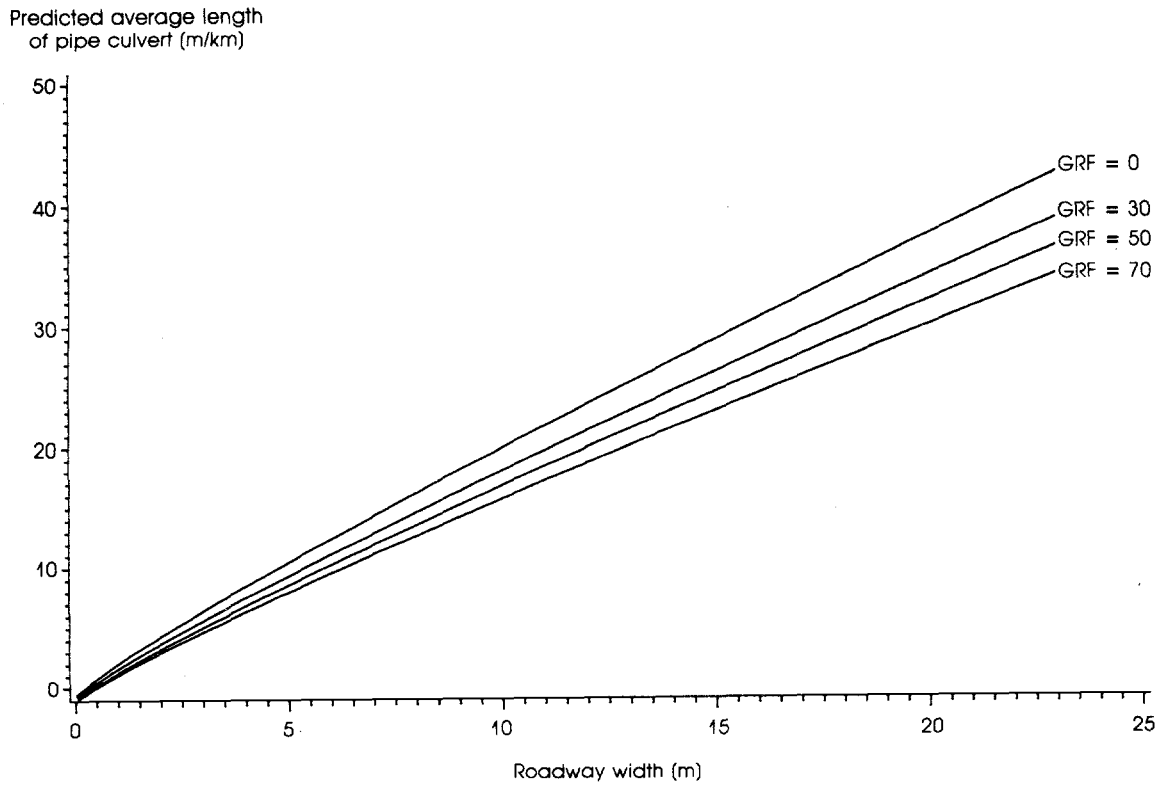
$$DRL = \begin{cases} 1.97\ [ALPC_{(after)} - ALPC_{(before)}] & \text{if } 0 \leq GRF < 10 \\ 1.74\ [ALPC_{(after)} - ALPC_{(before)}] & \text{if } 10 \leq GRF < 40 \\ 2.02\ [ALPC_{(after)} - ALPC_{(before)}] & \text{if } 40 \leq GRF \end{cases}$$

where $ALPC_{(after)}$ and $ALPC_{(before)}$ are computed using the above expression for ALPC based on the "before" and "after" values of the roadway width, respectively.

Box culverts.

$$ANBC = \begin{cases} 0.27 & \text{if } 0 \leq GRF < 10 \\ 0.15 & \text{if } 10 \leq GRF < 40 \\ 0.62 & \text{if } 40 \leq GRF \end{cases}$$

Figure 3.6: Predicted average pipe culvert length versus roadway width



Source: Based on analysis in Tsunokawa (1983). See also Aw (1981; 1982) and Markow and Aw (1983).

where variables and subscripts are as defined above. Although the number of box culverts is the same for new construction and widening, unit costs per culvert may be different and therefore should be reflected in the input data.

3.2.4 Bridges

New construction (new and old alignment)

The relationship for predicting the total floor area of small bridges (those having span lengths of less than 60 m) is given by:

$$AB = \begin{cases} 4.35\ RW & \text{if } 0 \leq GRF < 10 \\ 2.09\ RW & \text{if } 10 \leq GRF < 40 \\ 1.83\ RW & \text{if } 40 \leq GRF \end{cases}$$

where AB = the average floor area of bridges per unit length of road, in m^2/km.

Widening

No relationship is available.

APPENDIX 3A

FORMULATION AND ESTIMATION OF RELATIONSHIPS FOR PREDICTING ROAD CONSTRUCTION QUANTITIES

At the stage of highway sector planning and resource allocation where the range of investment options (e.g., with respect to the location and standards) to be examined are the widest, policy-makers need a method of construction cost prediction that requires minimal information inputs and yet produces cost estimates properly sensitive to a broad spectrum of design standards and terrain characteristics. These requirements are essential to enable cost trade-offs among road construction and maintenance and vehicle operating costs to be made at the sector level. After realizing that no such method existed in suitable form, the World Bank and MIT initiated in 1981 a small-scale collaborative study to develop a set of relationships for predicting road construction costs that would meet the broad requirements above. The first product as reported in Aw (1981, 1982) and Markow and Aw (1983) essentially consisted of a comprehensive data base and a set of preliminary relationships which, because of the heavy reliance on engineering principles in their formulation, represented a considerable improvement over their predecessors. These relationships were further refined into a form suitable for general applications, as reported in Tsunokawa (1983). This appendix provides a summary description of the data base and the development of the final relationships.

3A.1 Data Base

As detailed in Aw (1981) road construction data were compiled from 52 road projects located in 28 countries in Asia, Africa, and Central and South America. These countries include: Indonesia, New Guinea, the Phillipines, Taiwan and Thailand in the East Asia and the Pacific Region; Nepal and Pakistan in South Asia; Syria and Turkey in the Middle East; Ethiopia, Kenya, Malawi, Somalia, the Sudan, Swaziland, Uganda and Upper Volta in East and West Africa; Honduras, El Salvador and Panama in Central America; and Argentina, Bolivia, Chile, Columbia, Equador and Peru in South America.

The regions in which these road projects were constructed cover a broad spectrum of topographic, climatic and soil characteristics -- from flat plains in the Sudan to extremely mountainous areas in Nepal, from the abundance of monsoon rainfall in Pakistan to the dryness of inland Africa, and from areas of good soil materials to lands of poor road-making volcanic ash. Of the 236 observations (road sections) investigated, 42, 24 and 34 percent were in mountainous, rolling and flat areas, respectively. The types of construction varied from feeder roads to four-lane freeways, from earth roads to concrete paved roads, and from 30 to 100 km/h design speeds.

3A.2 Formulation and Estimation of Final Relationships

The original relationships developed by Aw and Markow provide

separate estimates for "flat," "rolling" and "mountainous" areas based on a satisfactory because although the predicted construction quantities are sensitive to terrain type, it is difficult to determine in borderline cases whether the terrain is "flat" as opposed to "rolling" or "rolling" as opposed to "mountainous". To overcome this problem, the revised relationships reported herein employ only a continuous description of terrain severity which can be measured on an objective basis. This objective description, called the ground rise plus fall and devised by Aw (1982), will be defined later in this appendix.

Several of the relationships formulated are non-linear in parameters and were estimated using a special non-linear statistical procedure (NLIN) provided in the commercially available Statistical Analaysis System (SAS) package. The following paragraphs provide a brief description of the model formulation and estimation results for the individual relationships.

Area of site clearance

The area of site clearance per unit road length (ACG) is hypothesized as a linear function of the roadway width in which the coefficients are, in turn, exponential functions of the ground rise plus fall. The statistical estimation using this model form yielded the following results:

$$\underset{}{ACG} = \underset{(0.8)}{1770} \exp \underset{(1.5)}{[0.0278\ GRF]} + \underset{(6.2)}{1610} \exp \underset{(1.2)}{[-0.0114\ GRF]}\ RW$$

$$R^2 = 0.51 \qquad \text{Number of observations} = 35$$

where the variables are as defined in the text and the figures in the parentheses are asymptotic t-statistics. The goodness of fit of this relationship is illustrated in Figure 3A.1 in which the observed values of the area of site clearance is plotted against the predicted.

As shown graphically in Figure 3.3, this relationship seems to indicate different practices in setting up the site clearance area under different terrain conditions. First, the sensitivity of the site clearance area to the roadway width is at its maximum for flat terrain (GRE = 0) and decreases as the terrain becomes more severe (GRE approaches 100 m/km). Second, for relatively narrow roads (RW = 5 m) the site clearance area increases with terrain severity, which is as expected; however, the reverse is indicated for relatively wide roads (RW = 25 m). There is no clear explanation for the latter observation, although there might be a greater tendency to economize on relatively expensive construction in mountainous terrain.

Earthwork volume

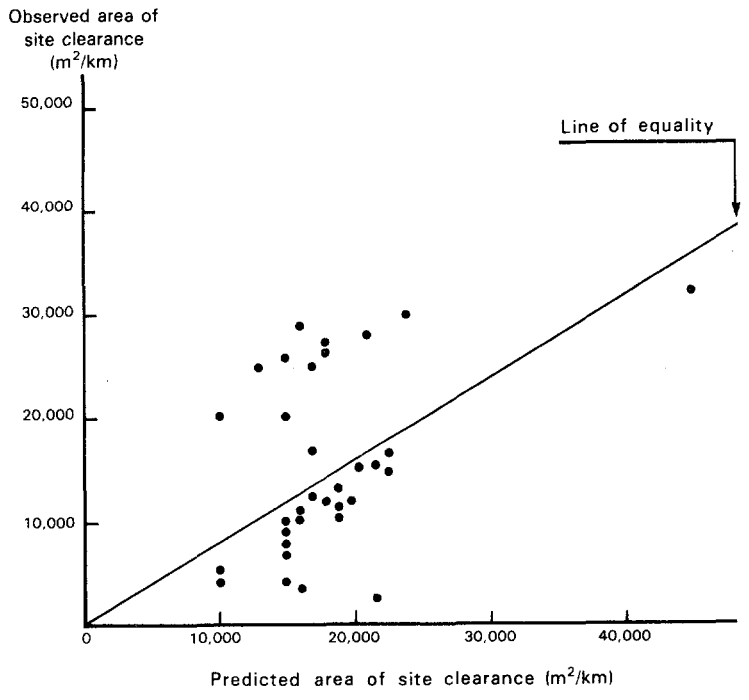
Among the formulas attempted for predicting the earthwork volume, the following was selected for use in HDM III:[1]

---

[1] This relationship is a revised version of the one reported in Tsunokawa (1983).

Figure 3A.1: Area of site clearance: observed versus predicted

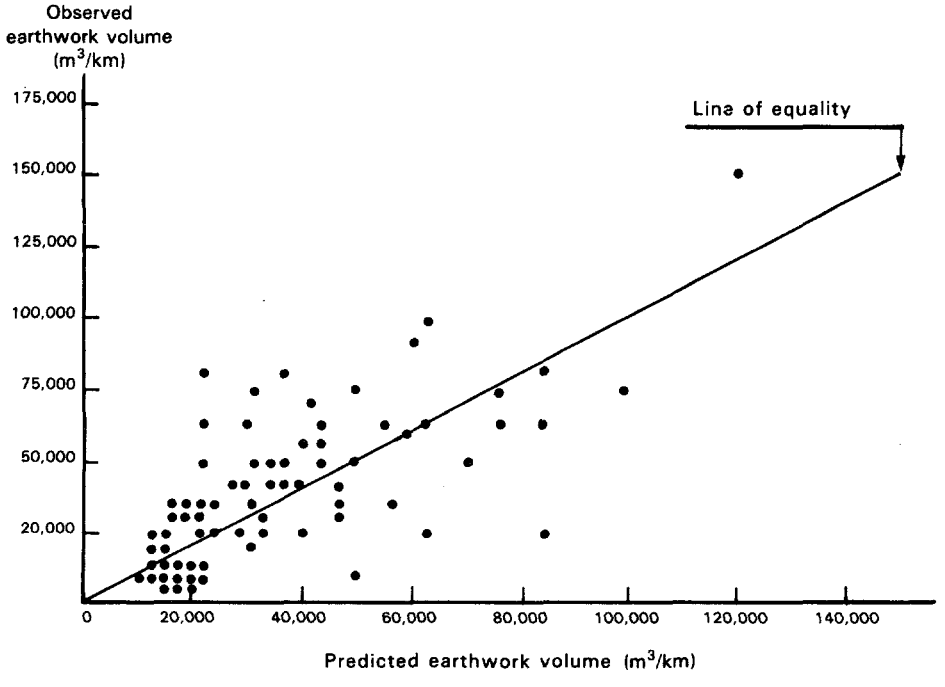
(half-page fig



Source: Based on analysis in Tsunokawa (1983). See also Aw (1981; 1982) and Markow and Aw (1983).

Figure 3A.2: Earthwork volume: observed versus predicted

(half pa



Source: Adapted from Tsunokawa (1983). See also Aw (1981; 1982) and Markow and Aw (1983).

where $H = 1.41 + 0.129\,G + 0.0139\,GRF$
$\quad\quad (7.5) \quad (3.7) \quad\quad (1.3)$
$G = GRF - RF$
$R^2 = 0.55$ Number of observations = 123

and the figures in the parentheses are asymptotic t-statistics. A plot of the observed against the predicted value of the earthwork volume is provided in Figure 3A.2.

This specification is based on the following simple physical construct of the road's cross-section. In a simple case of zero ground crossfall, which typically occurs in flat terrain, the cross-section of a fill section is assumed to be represented by a trapezoid, as shown schematically in Figure 3A.3(a). The volume of earthwork per unit road length (m^3/km) is given by:

$$EWV = 1000\,(RW + \cot {}_m\, H)\, H$$

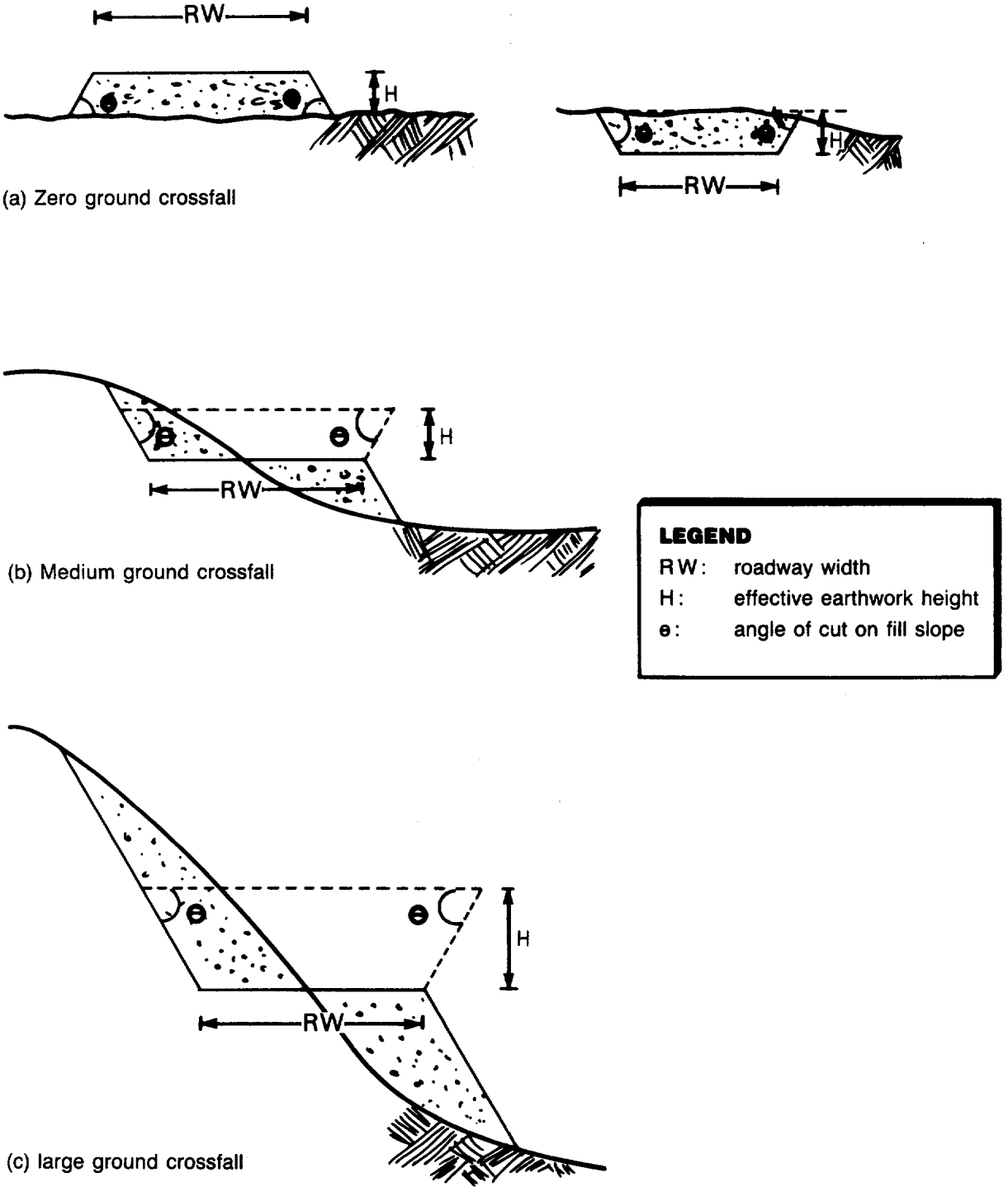
where ${}_m$ = the angle of the embankment slope, in radians;
H = the height of the embankment, in meters.

The same derivation applies to a cut section (see also Figure 3A.3) with the terms ${}_m$ and H now referring to the cut volume. In the case of medium and large ground crossfalls, which can be found in rolling and mountainous terrain, the volume of earthwork may also be represented by the same mathematical formula, but the physical construct is somewhat different. As depicted in Figures 3A.3(b) and 3A.3(c), the cross-sectional area of earthwork can be represented by an imaginary trapezoid with the shorter base equal to RW, but now with the base angle equal to the cut or fill slope and the height to an imaginary height. When this formula is supplied over an entire road stretch with cross-sections of varying shape and size, the term H is interpreted as the average of the heights of the real or imaginary trapezoids that represent these cross-sections. At this level of generalization, H is called the "effective earthwork height".

It is further hypothesized that the average effective height H is a function of terrain severity and the design standard of the road section; for simplicity, H is assumed to be a linear function of G and GRF. The expectation that the earthwork volume is relatively large for a road section of high standard relative to terrain severity (i.e., large G) is supported by the significant positive coefficient for G (0.129). The positive coefficient for GRF (0.0139), although not significant, is consistent with the expectation that the earthwork volume tends to increase with terrain severity for a given "relative" standard (G).

Although with small t-ratio (1.6), the estimated value of cot ${}_m$ of 0.731 appears to be reasonable. The values of G and GRF should be close to zero for flat terrain. When both G and GRF are zero, the predicted effective earthwork height, H, equals 1.41 meters. The height of a fill section in flat terrain is generally a function of the hydrologic and drainage conditions of the area. However, this value appears to be representative of the construction projects for which the data were

Figure 3A.3: Typical earthwork cross-sections and equivalent trapezoids



collected. In HDM-III, however, an option is provided to override the endogenously predicted value of H by a user-specified input when GRF is between 0 and 10 m/km.

Pipe culvert length

The relationship for predicting the aggregate length of pipe culvert was constructed as a product of two relationships, one for predicting the average length of a pipe culvert and another for predicting the number of pipe culverts per unit length of road. An attempt to derive a relationship which directly predicts the aggregate length of pipe culvert per unit road length was also made, but the former approach was preferred on the basis of better fit to the data.

Among the various specifications tested, the following was selected for predicting the average length of a pipe culvert in HDM-III:

$$\begin{aligned} ALPC &= 2.57 \exp(-0.00313\ GRF)\ RW^{0.895} \\ &\quad (5.20) \qquad (1.84) \\ R^2 &= 0.705 \qquad \text{Number of observations} = 75 \end{aligned} \tag{10.82}$$

where the figures in the parentheses are asymptotic t-statistics. As illustrated graphically in Figure 3A.4, ALPC is an increasing function of RW; this is because the exponent and the multiplicative term are both positive. The negative coefficient of GRF indicates that for a given roadway width (RW), the average pipe culvert length decreases as terrain severity (GRF) increases. A plausible explanation for this is the greater tendency to minimize construction cost in steep terrain.

Although a number of relationships were tested for predicting the number of pipe culverts per unit road length, no significant improvements were obtained over the following group averages:

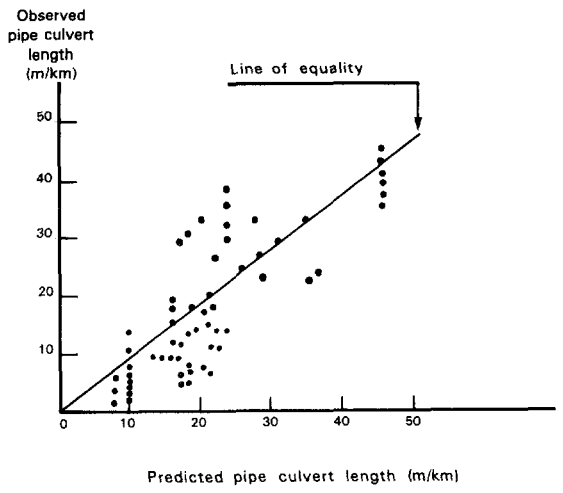
Average number of pipe culverts per km of road	Standard error of estimate	Number of observations	Range of ground rise + fall (m/km)
1.97	0.19	27	$0 \leq GRF < 10$
1.74	0.18	23	$10 \leq GRF < 40$
2.02	0.23	25	$40 \leq GRF < 100$

Source: Adapted from Aw (1981).

Thus the aggregate length of pipe culverts per unit length of road (m/km) is predicted in HDM-III as follows:

$$DRL = \begin{cases} 1.97\ ALPC & \text{if } 0 \leq GRF < 10 \\ 1.74\ ALPC & \text{if } 10 \leq GRF < 40 \\ 2.02\ ALPC & \text{if } 40 \leq GRF. \end{cases}$$

Figure 3A.4: Pipe culvert length: observed versus predicted



Source: Based on analysis in Tsunokawa (1983). See also Aw (1981; 1982) and Markow and Aw (1983).

Other quantities of road construction

Other quantities of the road construction, namely, the average numbers of box culverts (ANBC) and bridges (ANBR) per unit length of road are generally influenced by the terrain and climatic characteristics of specific locations. However, because of the lack of data for these explanatory variables, the quantities are predicted as group averages in HDM-III (shown with standard errors in parentheses):

		Range of ground rise plus fall (m/km)		
Variable	Description	$0 \leq GRF < 10$	$10 \leq GRF < 40$	$40 \leq GRF < 100$
ANBC	Number of small box culverts pert km (spans under 1.0 m)	0.27 (0.04)	0.72 (0.15)	0.62 (0.20)
ANBR	Number of small bridges per km (spans under 60 m)	0.217 (0.027)	0.104 (0.016)	0.091 (0.027)
Number of observations		43	26	16

Source: Adapted from Aw (1981).

Assuming that the average width of bridges is equal to the roadway width (RW), the average floor area of bridges per unit road length (m^2/km) is given by:

$$AB = 20\ ANBR$$

where 20 is the average span lenght of small bridges found in the data (in meters).

CHAPTER 4

Road Deterioration and Maintenance Submodel

The road deterioration and maintenance submodel performs the important function of linking construction standards (and costs), road maintenance standards (and costs), and road user costs through the road deterioration relationships. The HDM model considers these relationships in detail because the deterioration of road condition, ultimately manifest in the roughness of the road surface, causes significant increases in road user costs. The rate of deterioration and the effectiveness of maintenance together affect the timing, nature and costs of future investments in rehabilitation and the magnitude of returns in savings on user costs.

The submodel estimates the combined effects of traffic, environment and age on the condition of the road, given data on its construction and materials, and proceeds year by year to predict the change of surface condition under specified maintenance and rehabilitation policies throughout the course of the analysis period. Condition is predicted by distress mode; for paved roads: cracking, ravelling, potholes, rutting and roughness; for unpaved roads: roughness, material loss and passability. Parameters include the volume and loading of traffic, rainfall and moisture balance, initial road conditions, material strength properties and thicknesses, the variability of material behavior and construction quality, and a range of maintenance options. They cover paved roads with flexible pavements of bituminous surfacings and either granular or cemented base (but exclude portland cement concrete and thick asphalt base pavements), and unpaved roads of either gravel or earth surfacing. They apply mainly to engineered roads with adequate drainage, crossfall and foundation stability (below the subgrade), though some guidance is given on how to parameterize roads that do not conform in all respects. The models are applicable directly to non-freezing climates and with adjustments may apply in some freezing climates.

The relationships incorporated in this submodel are structured on mechanistic concepts of pavement behavior and are empirical models statistically-quantified (Paterson, 1987) from an extensive data base collected in the Brazil-UNDP study (GEIPOT, 1982). Some are probabilistic rather than deterministic in nature, predicting an array of outcomes. Some supplementary models from other sources are included and in minor models, where an adequate empirical base was lacking, engineering principles and judgment have been applied. The primary relationships have been verified on independent data sets (Paterson, 1987). Only a single set of relationships is included in the submodel but provision is made for the user to adapt these to local conditions where quantitative studies are available.

Road roughness throughout this chapter has been expressed in QI units for the sake of consistency with the rest of the model specification.

The relationships and all supporting documentation in Paterson (1987) are expressed in International Roughness Index (IRI) units. (Note: 1 m/km IRI = 13 counts/km QI_m.)

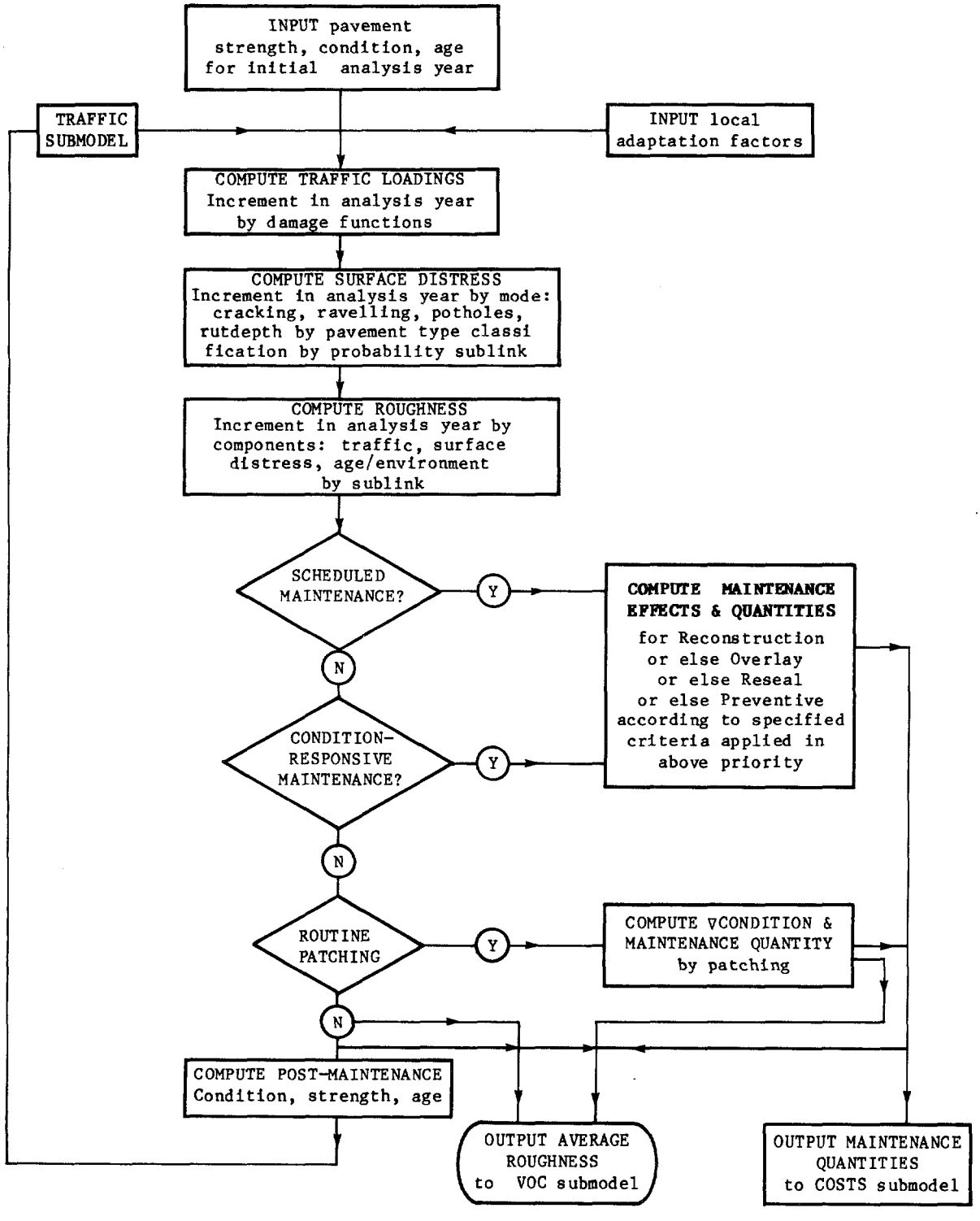
4.1 PAVED ROAD CONCEPTS AND LOGIC

4.1.1 General Concepts

In keeping with the structure of HDM III, road deterioration is computed as the incremental change in pavement condition during each analysis year due to traffic, environment and maintenance, and the current pavement condition is updated each successive year during the analysis period. This computational sequence is shown diagrammatically in Figure 4.1. The empirical relationships are compatible with this and are primarily recursive incremental models in which the predicted change in condition is a function of current condition, the cumulative traffic and rainfall, and various pavement parameters. This model form is able to accommodate pavements in any initial state of condition and age, facilitates the handling of environmental effects and interactive maintenance criteria, is particularly apposite to both validation and application in pavement management systems, and is appropriate to marginal cost evaluation. Historical data on trends of pavement condition are not a required input although where these are available they can be used indirectly to calibrate the relationships as discussed under local adaptation (Section 4.1.8).

Road pavements deteriorate over time under the combined effects of traffic and weather. Traffic axle loadings induce levels of stress and strain within the pavement layers which are functions of the stiffness and layer thicknesses of the materials and which under repeated loading cause the initiation of cracking through fatigue in bound materials and the deformation of all materials. Weathering causes bituminous surfacing materials to become brittle and thus more susceptible to cracking and to disintegration (which includes ravelling, spalling, and edge-breaking). Once initiated, cracking extends in area, increases in intensity (closer spacing) and increases in severity (or width of crack) to the point where spalling and ultimately potholes develop. Open cracks on the surface and poorly maintained drainage systems permit excess water to enter the pavement, hastening the process of disintegration, reducing the shear strength of unbound materials and thus increasing the rate of deformation under the stresses induced by traffic loading. The cumulative deformation throughout the pavement is manifested in the wheelpaths as ruts and more generally in the surface as an unevenness or distortion of profile termed roughness. Apart from, and interacting with, traffic, environmental effects of weather and seasonal changes cause distortions which also result in roughness. The roughness of a pavement is therefore the result of a chain of distress mechanisms and the combination of various modes of distress. Maintenance is usually intended to reduce the rate of deterioration, but certain forms such as patching may even increase the roughness slightly. Roughness is thus viewed as a composite distress,

Figure 4.1: Logic sequence of road deterioration and maintenance submodel: paved roads



Source: This study.

comprising components of deformation due to traffic loading and rut depth variation, surface defects from spalled cracking, potholes, and patching, and a combination of aging and environmental effects.

These concepts of pavement behavior are reflected in the interactive nature of the relationships for each distress mode and determine the sequence of computation shown in Figure 4.1. The pavement strength (which is summarized in an index, the modified structural number), the condition, and the age of the pavement at the beginning of the year are given, and the volume of traffic per lane is computed using two damage functions to reduce the spectrum of axle loading to a number of equivalent 80kN standard axle loads (ESA). The ages predicted for the initiation of cracking or ravelling vary with surface and base type, and when the current surfacing age exceeds those, the areas of cracking and ravelling progression are predicted. Potholing begins beyond a threshold of the area and severity of cracking and ravelling, and progresses by volume. The variability of material behavior is incorporated by estimating the expected times of early failures, median failures and late failures for equal thirds of the link length which are then treated as sublinks. The increments of rut depth and of roughness due to deterioration are then computed for each sublink. Local adaptation of the prediction relationships is effected through input "deterioration factors" which multiply the predicted values. The average values of roughness for each sublink over the year are output to the Vehicle Operating Cost submodel after routine patching maintenance, but before major maintenance, effects are computed.

Maintenance is applied at the end of the analysis year if the specified intervention criteria are met. Maintenance intervention is specified as either "scheduled" (i.e., at specific time intervals) or 'condition responsive' (i.e., at specified threshold levels of condition). There are five categories of maintenance, in order of descending priority: reconstruction, overlay, resealing, preventive treatment, and patching. The prioritization ensures, for example, that if the intervention criteria for both reconstruction and overlay were satisfied in a given analysis year then only reconstruction would be applied. Routine patching maintenance is therefore only applied when no major maintenance is applied in that year. It is regarded as having been applied continually during the year, unlike major maintenance which is applied at the end of an applicable year. The policy for each maintenance category specifies either the unit quantities to be applied, for example the thickness of overlay or reseal, the annual and maximum quantities of patching, etc., or the condition to be achieved, e.g., the strength after reconstruction. The quantities are neither endogenously "designed" nor optimized, and are computed solely on the basis of the specified criteria.

Following the application of maintenance, the maintenance quantities and costs are computed and output to the benefits-costs submodel. The values of the road condition parameters after maintenance (i.e., cracking, ravelling, potholing and patching areas, rut depth and roughness) are computed and become initial values for the next analysis year. The cycle continues through successive years to the final analysis year.

4.1.2 Computational Logic

Road deterioration is predicted through five separate distress modes, i.e., cracking, ravelling, potholing, rutting and roughness, as defined in Table 4.1 and illustrated in Figure 4.2. Surfacing distress, namely cracking, ravelling and potholing, is characterized by two phases, i.e.,

1. Initiation phase - the period before surfacing distress of a given mode or severity develops;

2. Progression phase - during which the area and severity of distress increases.

Separate sets of equations are provided for initiation and progression. In addition, cracking is characterized by two degrees of severity, "all cracking" (narrow and wide) and "wide cracking," which are subsequently combined in a composite index. Deformation distress, namely rut depth and roughness are continuous, and represented by only progression equations. As they are partly dependent upon the surfacing distress, they are computed after the change of surfacing distress in the analysis year has been computed.

As an aid to keeping track of the many variables used in this chapter, some conventions have been adopted as follows:

Subscripts

a [condition] at beginning of analysis year (after maintenance of the previous year);
b [condition] at end of analysis year (before maintenance);
d [change of condition] due to deterioration;
m [change of condition] due to maintenance.

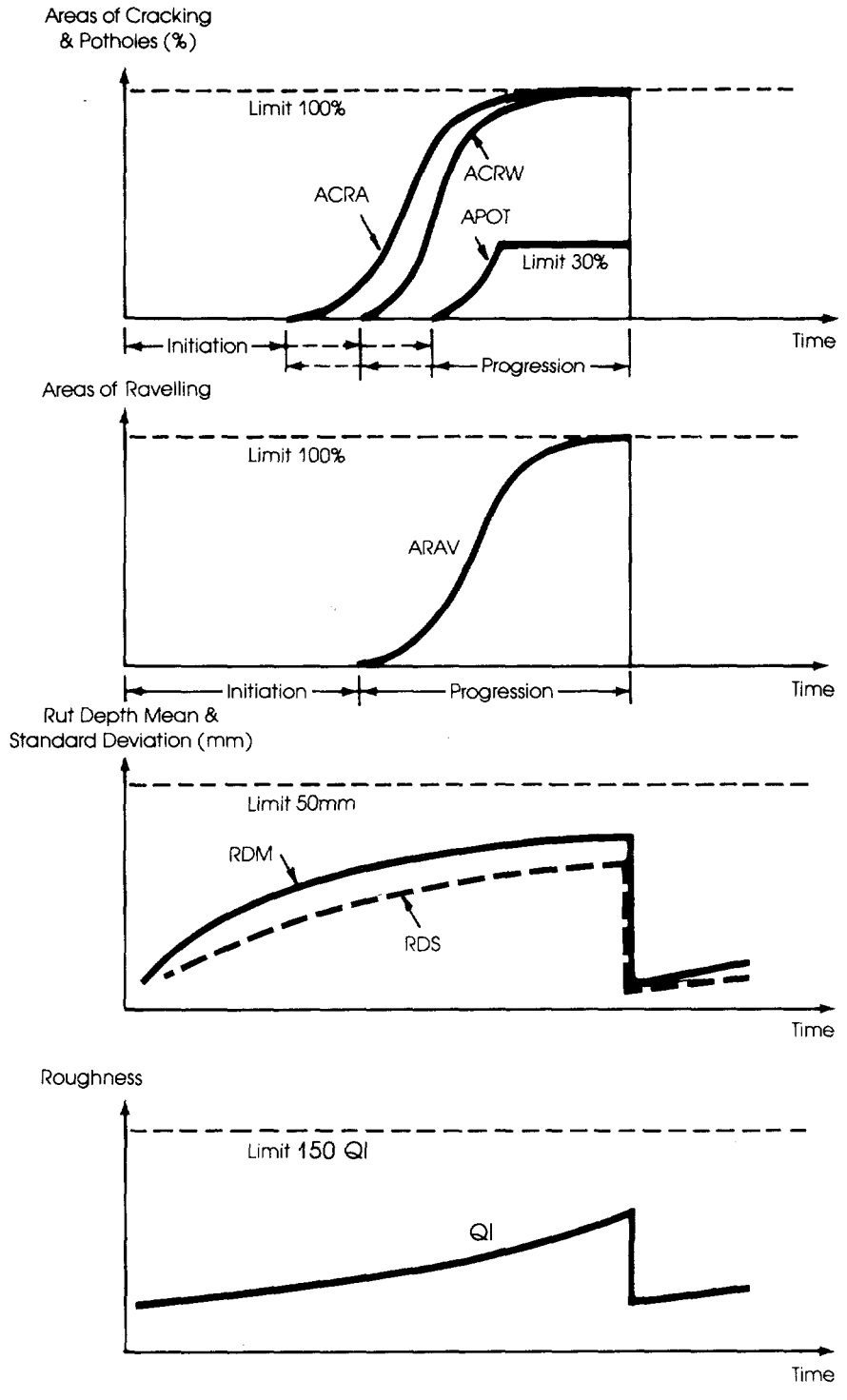
Prefixes

P previous value, before maintenance (e.g., PCRA in Table 4.2);
Δ change of condition variable during analysis year.

Also, in general, the letter A usually stands for "area" (except in the word "AGE"), T is for "time," and Y for "yearly" or "years." Other mnemonics will be apparent as the presentation progresses. The primary variables used from one analysis year to the next to define pavement condition, history and strength are defined in Table 4.2 and may be classified into groups as follows:

[CONDITION] = [ACRA, ACRW, ARAV, APOT, RDM, RDS, QI]
[HISTORY] = [AGE1, AGE2, AGE3]
[TRAFFIC] = [YE4, YAX]
[STRUCTURE] = [SNC, DEF, HS.., PCRA, PCRW, CRT, RRF]

Figure 4.2: Primary modes of pavement distress estimated in road deterioration and maintenance submodel



Source: This study.

Table 4.1: Definitions of distress measures

Measure		Definition
Area (of distress)	-	Sum of rectangular areas circumscribing manifest distress (line cracks are assigned a width dimension of 0.5m), expressed as percentage of carriageway area and section length;
All cracking	-	Narrow and wide cracking inclusive;
Narrow cracking	-	Interconnected or line cracks of 1-3 mm crack width (equivalent to AASHTO Class 2);
Wide cracking	-	Interconnected or line cracks of 3 mm crack width or greater, with spalling (equivalent to AASHTO Class 4);
Indexed cracking	-	Normalized sum of AASHTO Classes 2 to 4 cracking weighted by class, i.e.: $CRX = \sum_{i=1,4} ACR_i / \Sigma i$, and estimated by (Paterson 1987): $CRX = 0.62\ ACRA + 0.39\ ACRW$.
Ravelling	-	Loss of material from wearing surface;
Rut depth	-	Maximum depth under a 1.2 m straightedge placed transversely across a wheelpath;
Pothole	-	Open cavity in road surface with at least 150 mm diameter and at least 25 mm depth;
Roughness	-	Deviations of a surface from a true planar surface with characteristic dimensions that affect vehicle dynamics, ride quality, dynamic loads and drainage (ASTM E-867-82A) - typically in ranges of 0.1 to 100 m wavelengths and 1 to 100 mm amplitudes;
IRI	-	International Roughness Index, the reference measure expressing roughness as a dimensionless average rectified slope statistic of the longitudinal profile and defined in Sayers, Gillespie, and Paterson (1987);
QI	-	Quarter-car Index, roughness measure of the main Brazil-UNDP study data base, an average rectified slope statistic, where 13 counts/km $QI_m \simeq 1$ m/km IRI.

Source: This study.

Table 4.2: Definition of primary variables for pavement condition, history and strength

Variable	Definition
$ACRA_a$, $ACRA_b$	The total area of all cracking, in percent of the total carriageway area
$ACRW_a$, $ACRW_b$	The total area of wide cracking, in percent of the total carriageway area
$ARAV_a$, $ARAV_b$	The total area ravelled, in percent of the total carriageway area
$APOT_a$, $APOT_b$	The total area of potholing, in percent of the total carriageway area
AGE1	The preventive treatment age, defined as the time since the latest preventive treatment, reseal, overlay, reconstruction standard axle load and 520 kPa tire pressure) of the surfacing
AGE2	The surfacing age defined as the time since the latest reseal, overlay, reconstruction or new construction activity, in years
AGE3	The construction age, defined as the time since the latest overlay, reconstruction or new construction activity, in years
CMOD	The resilient modulus of soil cement, in GPa (required for cemented base pavements only)
COMP	The relative compaction in the base, subbase and selected subgrade layers, in percent (see note in 4.2.6)
CQ	The construction quality indicator for surfacing, where CQ = 1 if the surfacing has construction faults, = 0 otherwise
CRT	The cracking retardation time due to maintenance, in years (as elaborated in 4.3.3)
CRX_a, CRX_b	The total area of index cracking area, in percent of the total carriageway
DEF	The mean Benkelman Beam rebound deflection of the surfacing in both wheelpaths under 80kN standard axle load, 520 kPa tire presure, and 30°C average asphalt temperature), in millimeters
HBASE	The thickness of the base layer in the original pavement (required only for cemented base pavements), in millimeters
HSNEW	The thickness of the most recent surfacing, in millimeters
HSOLD	The total thickness of previous, underlying surfacing layers, in millimeters
MMP	Mean monthly precipitation, in m/month
PCRA	The area of all cracking before the latest reseal or overlay, in percent of the total carriageway area
PCRW	The area of wide cracking before the latest reseal or overlay, in percent of the total carriageway area
QI_a, QI_b	The road roughness, in QI units (see Figure 1.4 for conversion relationships, and note that the model accepts alternate units with linear conversion)
RDM_a, RDM_b	The mean rut depth in both wheel paths,in mm
RDS_a, RDS_b	The standard deviation of rut depth (across both wheel paths), in mm
RRF	The ravelling retardation factor due to maintenance (dimenssionless - as elaborated in 4.3.3)
RH	The rehabilitation indicator, where RH = 1 for surface types asphalt concrete overlay (OVSA) or open-graded cold-mix (OCMS) overlays, = 0 otherwise
SNC	Modified structural number of the pavement, computed as defined in Section 4.1.3.
YAX	The total number of axles of all vehicle classes for the analysis year, in millions/lane
YE4	The number of equivalent standard axle loads for the analysis year based on an axle load equivalency exponent of 4.0, in millions/lane

Source: This study.

The computational logic has the following sequence (see also Figure 4.1):

1. The pavement condition at the beginning of the analysis year is initialized either from input data if it is the first year of the analysis or the first year after construction, or otherwise from the result of the previous year's condition after maintenance (see Section 4.2.1):

   $[CONDITION]_a$ = initialized, or updated after maintenance of previous cycle (see (v) below).
   $[HISTORY]_a$ = $[HISTORY]_b + 1$
   $[TRAFFIC]_a$ = updated from traffic submodel
   [STRUCTURE] = updated as result of previous maintenance

2. The surface conditions before maintenance at the end of the year are predicted (see Sections 4.2.2-4.2.5):

   $$[ACRA, ACRW, ARAV, APOT]_b = [ACRA, ACRW, ARAV, APOT]_a + \Delta[ACRA, ACRW, ARAV, APOT]_d$$

3. The rut depth and roughness conditions before maintenance at the end of the years are predicted (see Sections 4.2.6-4.2.7):

   $$[RDM, RDS, QI]_b = [RDM, RDS, QI]_a + \Delta[RDM, RDS, QI]_d$$

4. Maintenance intervention criteria are applied to determine the nature of maintenance to be applied, if any (see Sections 4.3.3-4.3.7),

   Condition responsive:

   if $[ACRW, ARAV, APOT, QI]_b >= [ACRW, ARAV, APOT, QI]_{intervention}$

   or Scheduled:

   if $[AGE1, AGE2, AGE3]_b >= [AGE1, AGE2, AGE3]_{intervention}$

5. Highest-ranking applicable maintenance is applied and the effects on pavement condition computed (see Sections 4.3.3-4.3.7):

   $$[CONDITION]_{a(next\ year)} = [CONDITION]_b + \Delta[CONDITION]_m$$

4.1.3 Pavement Structural Characteristics

In performance prediction models it is necessary to use measures of pavement strength which summarize the complex interactions between material types and stiffnesses, layer thicknesses and depths, subgrade stiffness and surface condition. Measures which have proved effective or popular in the past have included, <u>inter alia</u>:

Structural number (AASHO) with soil support value (AASHTO) or modified structural number (TRRL);
Surface deflection under creep loading (Benkelman Beam);
Surface deflection and curvature under dynamic cyclic loading (Dynaflect, Road Rater, etc.); and
Surface deflection and curvature under dynamic impact loading (Falling Weight Deflectometer).

In the empirical research (Paterson 1986b, 1987), the modified structural number (SNC) was found to be the most statistically significant measure of pavement strength affecting the deterioration of pavements, and is thus the primary strength parameter in the prediction relationships. Surface peak deflection measurements under standard loadings were generally weak predictors of performance without supplementary strength parameters, though Benkelman Beam deflection (DEF) enters some relationships. A fair degree of correlation exists amongst strength and deflection measures however, and the user may estimate SNC or DEF for input, either directly or from alternative measures as follows.

The modified structural number, SNC, is defined as a linear combination of the layer strength coefficients a_i and thicknesses H_i of the individual layers above the subgrade, and a contribution from the subgrade denoted by SNSG, i.e.:

$$SNC = 0.0394 \sum_{i=1}^{n} a_i H_i + SNSG$$

where a_i = the strength coefficient of the ith layer as defined in Table 4.3 and shown in Figure 4.3(a), (b);

H_i = the thickness of the ith layer provided that the sum of thicknesses, H_i is not greater than 700 mm, in mm;

n = the number of pavement layers;

SNSG = the modified structural number contribution of the subgrade, given by:

$$SNSG = 3.51 \log_{10} CBR - 0.85 (\log_{10} CBR)^2 - 1.43; \text{ and}$$

CBR = the California Bearing Ratio of the subgrade at <u>in situ</u> conditions of moisture and density, in percent.

The strength variables are updated following maintenance to take account of changes in pavement strength due to overlaying, resealing and cracking. The models are sensitive to these variables, and input values should be selected carefully. The user is required to specify only one of the primary variables, SNC or DEF, but predictions can be improved by specifying both values. When only one value is supplied, the other

Table 4.3: Pavement layer strength coefficients

Pavement layer	Strength coefficient a_i	
Surface course		
Surface treatments	0.20 to 0.40	
Asphalt mixtures (cold or hot premix of low stability)	0.20	
Asphalt concrete (hot premix of high stability)[1]		
MR_{30} = 1500 MPa	0.30	
MR_{30} = 2500 MPa	0.40	
MR_{30} = 4000 MPa or greater	0.45	
Base course		
	Max. axle load > 80kN	Max. axle load < 80kN
Granular materials [2]		
CBR = 30% [3]	0.	0.07
CBR = 50%	0.	0.10
CBR = 70%	0.10	0.12
CBR = 90%	0.12	0.13
CBR = 110%	0.14	0.14
Cemented materials [4]		
UCS = 0.7 MPa		0.10
UCS = 2.0 MPa		0.15
UCS = 3.5 MPa		0.20
UCS = 5.0 MPa		0.24
Bituminous materials [5]		0.32
Subbase course and selected subgrade layers (to total pavement depth of 700 mm)		
Granular Materials [6]		
CBR = 5%		0.06
CBR = 15%		0.09
CBR = 25%		0.10
CBR = 50%		0.12
CBR = 100%		0.14
Cemented materials		
UCS > 0.7 MPa		0.14

(See footnotes on next page.)

[1] Applicable only when thickness $H_1 \geq 30$ mm. MR_{30} = resilient modulus by indirect tensile test at 30° C.

[2] $a_i = (29.14\ CBR - 0.1977\ CBR^2 + 0.00045\ CBR^3)\ 10^{-4}$; the coefficient a_i may be increased by 60 percent if CBR > 70 and the subbase is cement- or lime-treated. Note: $a_i = 0$ for CBR < 60 when maximum axle loading exceeds 80 kN.

[3] CBR = California Bearing Ratio (in percent) determined at the equilibrium in situ conditions of moisture content and density.

[4] $a_i = 0.075 + 0.039\ UCS - 0.00088\ UCS^2$; where UCS = Unconfined Compressive Strength in MPa at 14 days. 'Cemented' implies development of tensile strength through portland cement or lime-treatment, or the use of certain flyash, slag, lateritic or ferricrete materials that are self-cementing over time.

[5] Dense-graded bitumen-treated base of high stiffness, e.g., MR_{20} = 4000 MPa, resilient modulus by indirect tensile test at 20°C.

[6] $a_i = 0.01 + 0.065 \log_{10} CBR$.

Source: Adaptation of TRRL Report LR673 (Hodges et al., 1975), TRDF Final Report V Brazil (GEIPOT, 1982), and NITRR manual TRH4, 1978, South Africa.

variable value is computed endogenously by the following approximate relationships (Paterson, 1987) shown in Figure 4.3(c):

$$DEF = \begin{cases} 6.5\ SNC^{-1.6} & \text{if base is not cemented} \\ 3.5\ SNC^{-1.6} & \text{if base is cemented} \end{cases}$$

$$SNC = \begin{cases} 3.2\ DEF^{-0.63} & \text{if base is not cemented} \\ 2.2\ DEF^{-0.63} & \text{if base is cemented.} \end{cases}$$

Note: The standard error (S.E.) and valid range of these estimations are:

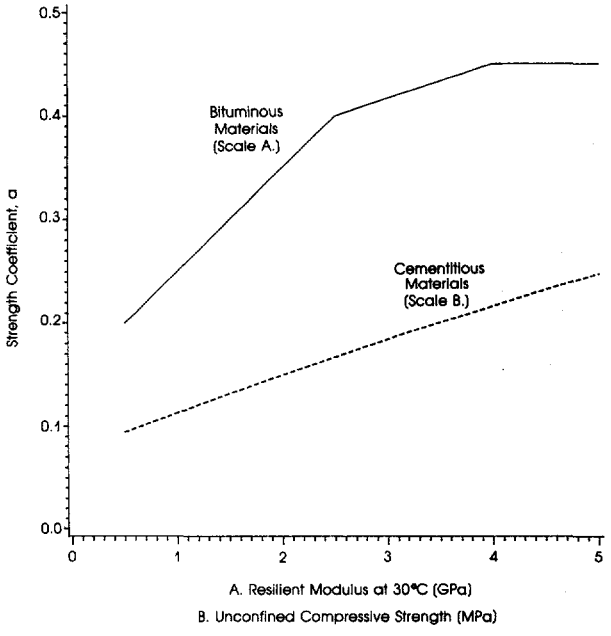
for DEF: S.E. = 0.34 mm; range = 0.13 to 2.0 mm; and

for SNC: S.E. = 1.24; range = 1.5 to 7.7.

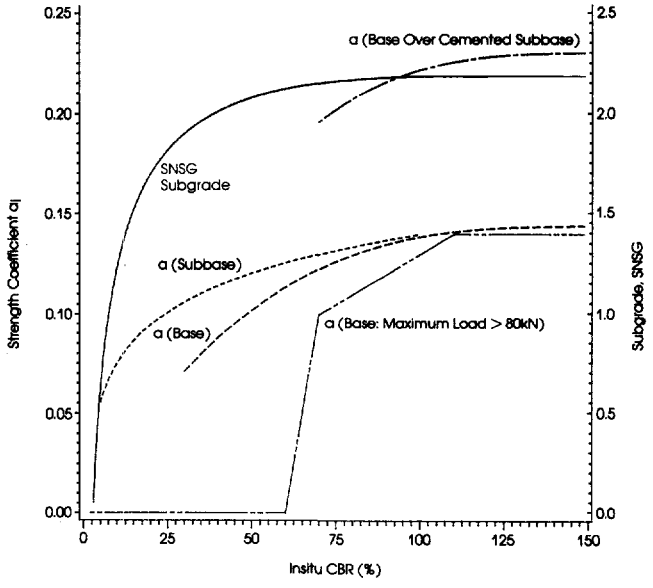
The alternative, dynamic, deflection measures may be used to estimate these inputs, but the relationships depend upon pavement configuration as discussed in Paterson (1987), and the user may need to generate a specific conversion appropriate to the circumstances.

Figure 4.3: Charts for estimating pavement strength and deflection parameters values

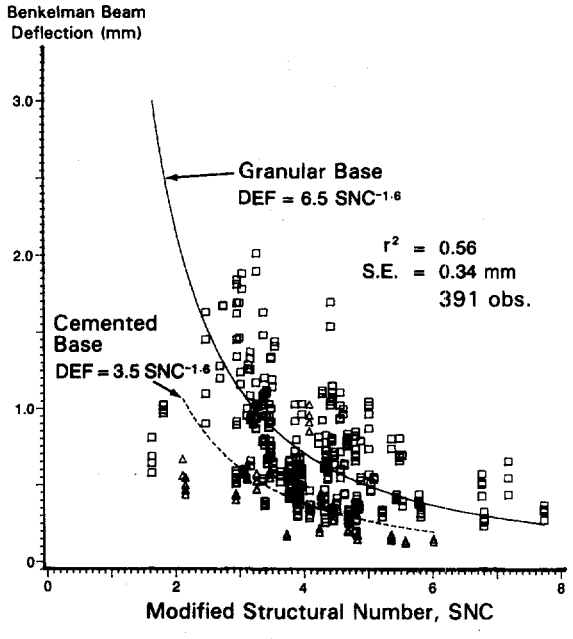
(a) Layer strength coefficients - bituminous and cemented materials



(b) Layer strength coefficients for unbound materials and estimation of subgrade support



(c) Estimation of Benkelman Beam Deflection (DEF) or Modified Structural Numbers (SNC)



Source: This study.

4.1.4 Pavement Classification

For the prediction of surfacing distress (cracking, ravelling and potholing), neither SNC nor DEF are sufficient predictors for all pavement types. Thus, pavements are classified by seven surface types and three base types, totalling twenty different pavement categories (one combination is not used in practice). Within each category, the greater number of surfacing and base combinations used in practice have statistically similar behavior, distinguished otherwise through the explanatory variables appearing in a specific distress equation. The seven surface types are:

1. Surface treatment (ST);
2. Asphalt concrete (AC);
3. Slurry on surface treatment (SSST);
4. Reseal on surface treatment (RSST);
5. Reseal on asphalt concrete (RSAC);
6. Open graded cold mix on surface treatment (OCMS); and
7. Asphalt overlay or slurry seal on asphalt concrete, and asphalt overlay on surface treatment (OVSA).

Two of the surface types, surface treatment and asphalt concrete, apply to original, new or reconstructed pavements. The other types define surfaces after a full-width maintenance treatment or rehabilitation of an existing pavement, as illustrated in Table 4.4. Slurry on surface treatment and cold mix on surface treatment are types which may be either original surfacings or the result of certain maintenance operations (resealing with slurry and overlaying with cold mix, respectively). The other surface types result from other combinations of reseals, overlays and original surfacings. Note that the categorization groups together the combinations having similar behavior in order to restrict the number of categories.

The three base types are: granular, cement stabilized and bituminous. These base types are established as originally constructed or by pavement reconstruction and assumed to be unchanged throughout the pavement life unless reconstruction takes place. In the classification used herein, pavement reconstruction includes all major rehabilitation that changes the classification or strength of the base.

4.1.5 Traffic Characteristics

For paved roads, traffic is defined by two variables which represent both the volume and loading aspects of mixed vehicular traffic. These are the flow of all vehicle axles (YAX) and the flow of equivalent 80 kN standard axle loads (YE4), both expressed on an annual, millions per lane basis as defined in Table 4.2.

In HDM-III, the road link or subsection comprises the full road width. Directional symmetry is assumed in the traffic flow and road deterioration submodels so that the variables in this submodel express average pavement condition, average traffic flow and average equivalent axle loading across lanes and directions.

Table 4.4: Surface type classification after maintenance treatment

Classification before maintenance	Classification after indicated maintenance			
	Surface treat-ment(RS)	Slurry seal (SS)	Cold mix overlay (CM)	Asphalt concrete (OV)
Surface treatment (ST)	RSST	SSST	OCMS	OVSA
Slurry on surface treatment (SSST)	RSST	SSST	OCMS	OVSA
Reseal on surface treatment (RSST)	RSST	SSST	OCMS	OVSA
Cold mix on surface treat-ment (CMST)	RSAC	SSST	OCMS	OVSA
Asphalt concrete (AC)	RSAC	OVSA	OCMS	OVSA
Reseal on asphalt concrete (RSAC)	RSST	OVSA	OCMS	OVSA
Asphalt overlay on asphalt concrete (OVAC)	RSAC	OVSA	OCMS	OVSA

Note: Only surface types ST, SSST, OCMS and AC can be treated as new construction in the model.
Source: This study.

To compute YAX and YE4, the total numbers of vehicle axles and equivalent standard axles applicable during the analysis year are determined in the traffic submodel (Chapter 2) and divided by the effective number of lanes, denoted by ELANES. The ELANES variable may either be specified by the user or take a default value expressed as a function of road width, i.e.:

$$\text{ELANES} = \begin{cases} 1.0 & \text{if } W < 4.5 \\ 1.5 & \text{if } 4.5 < W < 6.0 \\ 2.0 & \text{if } 6.0 < W < 8.0 \\ 3.0 & \text{if } 8.0 < W < 11.0 \\ 4.0 & \text{if } 11.0 < W \end{cases}$$

where W = carriageway width, in meters.

4.1.6 Environment and Geometry

Environmental factors are incorporated in the deterioration models through two variables, the modified structural number (SNC) and the mean monthly precipitation (MMP). The effects of aging and weathering are represented through time variables incorporated in the prediction models.

The modified structural number (SNC) includes the contribution of the subgrade (SNSG) and the strength coefficients of the material properties (a_1) under the in situ conditions of the local environment of the pavement. The combined effects of drainage and rainfall on the behavior of the pavement are thus represented through the effects of moisture on the strength properties and stiffness (e.g. resilient modulus) of the materials in each pavement layer. An increase in the moisture content of a material above the optimum associated with its density causes a decrease in the shear strength, and often a decrease in the stiffness, of the material. Thus it usually causes a decrease of the modified structural number and increase of the deflection.

Such increases in the moisture content of a pavement layer above its optimum can be caused by poor drainage, the rise of a shallow watertable due to seasonal increases in rainfall, or the ingress of water through either cracking in the pavement or ponding on the shoulders. Such effects are particularly noticeable in cuttings, and on grades where water entering the pavement in the upper part courses down the grade along the layers.

The equilibrium moisture content of a pavement material, which is the field moisture content when the above transient influences are not present, is a function of the soil moisture suction, and is influenced by the depth of watertable, soil type and evapotranspiration potential. The equilibrium moisture content (EMC) in pavement layers can be estimated using the following relationship (Paterson, 1987):

$$EMC = 0.26\ P075 + 0.019\ IM + 2.10$$

where EMC = the equilibrium moisture content, in percent by mass;
P075 = the amount of material passing 0.075 mm sieve, in percent by mass; and
IM = Thornthwaite's moisture index, in the 1955 classification where $-100 \leq IM \leq +100$.

When the watertable is stable but shallow, i.e., within the region of influence of the pavement layers, the moisture content will be higher than estimated by the above. Only under conditions where the material properties change significantly with season, for example a fluctuating shallow or perched watertable or poor drainage, is it necessary to estimate different seasonal states (one of which may represent a saturated state). Under these circumstances, the structural number should be an appropriate weighting of the seasonal states as follows (based on the contribution to roughness progression):

$$SNC = \frac{SNC_a \; SNC_b}{\left[a \; SNC_b^5 \; + b \; SNC_a^5 \right]^{0.2}}$$

where SNC_a = the modified structural number applying for season "a";
SNC_b = the modified structural number applying for season "b";
a, b = the duration of the applicable season (fraction of 1 year) where a + b = 1.

If the seasonal values of structural number cannot be determined directly from material properties, then use can be made of the seasonal deflections to estimate the seasonal structural number as described in Section 4.1.3.

In addition to the above general effects of climate and drainage on pavement behavior through the structural number, water entering the pavement through a permeable or cracked surface can cause significant shallow-seated distress in the base and surfacing. This effect is evident though the rainfall and cracking parameters in the rutting predictions (Section 4.2.6), though it may be understated for a poorly-drained base or highly moisture-susceptible base material. In other modes of distress, the effect was not statistically significant in the Brazil-UNDP empirical base (Paterson, 1987). As this may have been due to the high intensity, short duration nature of tropical rainfall, fairly high standards of pavement crossfall and drainage, and moderately free-draining soils, some adjustments may need to be made in respect of adverse conditions, e.g., poor drainage, suscrptible soils and high rainfall in subtropical and temperate regions, through the deterioration factors (Section 4.1.8). For example, the factor for cracking progression, K_{cp} (but not for cracking initiation), and possibly for rut depth, K_{rp}, may need to be increased.

The environmental-age component in the roughness prediction model takes acount of more general climatic effects, and provisional values for the coefficient have been established through cross-validation of the model on data in various climates (Paterson, 1987), as given in Table 4.5. The coefficient quantifies the time-dependent proportional change in roughness, and the user adjusts for this through the deterioration factor, K_{ge} (Section 4.1.8).

No geometric effects on deterioration are present. Gradient was not found to have significant effects in the empirical study but can give rise to drainage problems on short sections. Road width effects have not yet been studied sufficiently and an adaptation of the submodel to include them is under development (see Hoban, 1987).

4.1.7 Variability and Uncertainty

Road segments which have identical mean values of pavement strength indicators such as modified structural number, surface type, etc., are often observed to deteriorate at different rates. Structural

properties, drainage characteristics and construction quality vary along the road length and between roads of the same type. Uncertainty in the predictions of a statistical model arise through the limitations of the variables that can be practically included and the inference space of the model's derivation.

The model automatically divides the road section (part of a link) into three "subsections" of equal length, identified as (relatively) "weak," "medium," and "strong" before starting simulation. Although all the three subsections are assumed to have identical nominal pavement characteristics, the "weak" subsection deteriorates the fastest and the "strong" subsection the slowest. This assumption is expressed in the HDM in terms of the occurrence distribution factor, F, appearing in each relationship to predict the time at which cracking or ravelling starts. That is, these subsections employ the same basic distress initiation models but with occurrence distribution factors to account for the different rates of deterioration. The factors are not applied to any distress progression models. The values of the occurrence distribution factors used in the HDM model are predetermined and represent equal thirds of the probability distribution of failure based on the statistical analysis conducted in the Brazil-UNDP study (Paterson, 1987). They are listed in subsequent Sections (4.2.2-4.2.3).

4.1.8 Local Adaptation (Deterioration Factors)

It is not practical to include certain material properties and environmental conditions which may be specific to a region or country in the road deterioration relationships. These include, for example, the source and type of binder or aggregate, rainfall intensity, level of solar radiation, specific construction practices, equipment or material specifications, etc. The user may take account of these effects for specific sets of links by specifying a "deterioration factor," which is a linear multiplier of the prediction, for each mode of distress on the input form. In addition to the deterioration factor, an integer (CQ) representing the construction quality of surface treatments (0 = good, 1 = poor) has been provided to permit the user to evaluate the value of quality control and effectiveness of construction specifications.

Local adaptation of the HDM relationships through the exogenous specification of the deterioration factor should preferably be based on balanced, quantitative studies of specific distress functions under local conditions. Care should be taken to avoid the bias often associated with small samples. A representative sample would include a range of pavement age from young to old, a range of traffic volumes from low to high, and (when possible) a range of pavement strength within a given traffic volume and loading class. Collinearity between age and traffic volume should also be avoided (e.g., low volume, old age; high volume, young age). Where such a study is not available, use can be made of average representative values of performance indicators that might be determined. For example, the expected life of a surface treatment before resealing may be 12 years in the local region but the prediction of the HDM model for the same volume of traffic is 9 years, which suggests a value of the deterioration factor, $K_{ci} = 12/9 = 1.3$.

Table 4.5: Recommended values of environmental coefficient 'm' in roughness progression model, for various classes of climate

Moisture classification	Moisture index[1]	Temperature classification[2]		
		Tropical nonfreezing	Subtropical nonfreezing	Temperate freezing
Arid	-100 to 61	0.005	0.010	0.025
Semiarid	-60 to -21	0.010	0.016	0.035
Subhumid	-20 to +19	0.023	0.030	0.050
Humid	20 to 100	0.030	0.040	0.070+

Note: In the HDM-III model, $m = 0.023\ K_{ge}$.
[1] After Thornthwaite (1955).
[2] Definition of these classes is as yet uncertain: Tropical means warm temperatures 15 to 40°C and small range; Subtropical includes warm, high range (5 to 60°C) and cool, moderate range (-10 to 30°C) temperatures; Temperate freezing includes climates with annual pavement freezing (this last class may require either subdivision or reformulation of the model).
Source: Paterson (1987).

It is expected that the cracking, ravelling and potholing models are the most likely to require local adaptation. For example, in regions other than Brazil, the predicted life before cracking initiation for asphalt concrete may need to be lengthened by specifying, say, K_{ci} = 1.2 to 1.5, and likewise the predicted rate of cracking progression may have to be retarded by specifying, say, K_{cp} = 0.7 to 1.0. The predictors for other surfacing types will usually need less adjustment. Potholing progression may need suppression for good quality base materials, dense-graded with good cohesion, e.g., K_{pp} = 0.2 (see 4.2.4). The factor K_{ge} for the environmental roughness-age term in the roughness prediction model, should be derived as a ratio of the coefficient "m" from Table 4.5, as follows:

$$K_{ge} = m/0.023$$

4.2 PAVED ROAD DETERIORATION PREDICTION

4.2.1 Variables at the Beginning of the Analysis Year

At the beginning of the analysis year the traffic loading variables (YAX and YE4) are computed for the year based on the user-specified traffic data (Series E) and vehicle axle load characteristics (Series D).

The values of the variables that describe the environment (MMP), road geometry (ELANES and W), pavement characteristics (SN, SNC, SNSG,

DEF, HSNEW, HSOLD, HBASE, CQ, RH, CMOD, COMP) are provided in one of three ways:

1. When the analysis year is neither the first year of the analysis period nor a construction opening year, defined as the year immediately following the effective completion year of a construction option affecting the section, values are provided from the preceding analysis year;

2. When the analysis year is the first year of the analysis period, values are provided from the input data of existing link characteristics (Series A); or

3. When the analysis year is a construction opening year, values are provided from the input data of construction characteristics (Series B).

The variables that describe <u>pavement history</u> (AGE1, AGE2 and AGE3) and <u>pavement condition</u> ($ACRA_a$, $ACRW_a$, $ARAV_a$, $APOT_a$, RDM_a, RDS_a, QI_a, CRT, RRF, PCRA and PCRW) are initialized in different manners as follows:

1. When the analysis year is neither a construction opening year nor the first analysis year, values are provided from the preceding year, i.e.,

   $[AGE1, AGE2, AGE3]_{after} = [AGE1, AGE2, AGE3]_{before} + 1$

   $[CONDITION]_a = [CONDITION]_{after\ maintenance\ of\ previous\ year}$

2. When the analysis year is the first year of the analysis period, values are provided from the input data of existing link characteristics (Series A);

   $[AGE1, AGE2, AGE3]_{after} = [AGE1, AGE2, AGE3]_{input} + 1;$

   $[CONDITION]_{after} = [CONDITION]_{input}$

3. When the analysis year is a construction opening year, the variables are initialized to reflect a new pavement, as follows:

   $[AGE1, AGE2, AGE3] = 1$

   $[ACRA_a, ACRW_a, ARAV_a, APOT_a] = 0$

   $[RDM_a, RDS_a, PCRA, PCRW] = 0$

   $QI_a = QII_o$

   $CRT = 0, RRF = 1$

where QII_o = the initial road roughness after pavement construction or reconstruction specified by the user or

provided as a default value according to the surface types shown below (note that surface type RSAC is not permitted in the construction option data):

25 QI for AC or OVSA;
35 QI for ST, SSST and RSST; and
45 QI for OCMS.

4.2.2 Cracking Initiation and Progression

Cracking initiation and progression are predicted for two classes, all cracking and wide cracking (defined in Table 4.1), through separate sets of relationships. These are assigned to different surfacing and base types following the classification of Section 4.1.4, as shown in Table 4.5. These are based on empirical results for each type, with the exception of the cemented-base maintenance surfacings and all bituminous base types which have been assigned the most appropriate relation. The cracking predominantly comprises crocodile and irregular cracking related to traffic and oxidation effects. For cemented base pavements, the predictions include linear cracking from shrinkage effects. However linear cracking from subsidence and edge moisture movements, and cold temperature cracking, are not included.

The uncertainties involved in predicting cracking initiation are represented by the values of the occurrence distribution factors F_c in Table 4.6. The factors multiply the mean predictions and yield the ages at which the cracking of the three subsections of each link representing early, median and late failures (or weak, medium, strong) are expected.

Cracking initiation

Initiation is defined by distress extending over 0.5 percent of the subsection area. The meaning of the variables used is illustrated in Figure 4.4. The relationships for predicting the time to initiation (which is the surfacing age) for all cracking (TYCRA) and wide cracking (TYCRW) are given in Tables 4.8 and 4.9. The supplementary variables used here are defined as follows:

TYCRA = the predicted number of years to the initiation of narrow cracks since last surfacing or resurfacing (when the surfacing age AGE2 = 0);

TYCRW = the predicted number of years to the initiation of wide cracks since last surfacing or resurfacing (when the surfacing age AGE2 = 0).

K_{ci} = the user-specified deterioration factor for cracking initiation (default value = 1);

F_c = the occurrence distribution factor for cracking initiation for the subsection (the values used in HDM-III are listed in Table 4.5);

Table 4.6: Cracking prediction relationships by pavement type

Surface type	Base type		
	Granular	Cemented	Bituminous
Surface treatment (ST)	A	B	-
Asphalt concrete (AC)	C	B	C
Slurry on surface treatment (SSST)	D	F [1]	G
Reseal on surface treatment (RSST)	E	F, H [2]	H
Reseal on asphalt concrete (RSAC)	H	F, H [2]	H
Cold mix on surface treatment (OCMS)	D	F	G
Asphalt overlay on asphalt concrete on surface treatment (OVSA)	G	F	G

Notes: Pavement type is defined by surface type and base type. The characters in the table represent the relationships described in the text to be employed for each of the pavement types.
[1] In HDM-III (1987), relationship F has been disabled and replaced by relationship B.
[2] Relationship F for initiation and H for progression.
Source: This study.

Table 4.7: Values of the occurrence distribution factors, F_c for cracking initiation relationships, applying to one-third subsections of each section.

Relationships	Subsection		
	Weak	Medium	Strong
A, D, E	0.55	0.98	1.48
B, F	0.74	1.01	1.25
C, G, H	0.51	0.96	1.53

Source: Paterson (1987).

HSE = the effective thickness of the surfacing layers defined as HSE = min [100; HSNEW + (1 - KW) HSOLD] (see Figure 4.4);

KW = a variable for indicating the presence of wide cracking in the old surfacing layers, defined as KW = min [0.05 max (PCRW - 10; 0); 1]; and

KA = a variable for indicating the presence of all cracking in the old surfacing layers defined as KA = min [0.05 max (PCRA - 10; 0); 1].

The relationships for cracking initiation in original asphalt concrete and double surface treatment surfacings of granular base pavements, and all surfacings of cemented-base pavements are illustrated in Figures 4.5 and 4.6 respectively. The curves in Figures 4.5 and 4.6(a) between the time to cracking and the equivalent axle load flow, show an interaction between aging and traffic effects. At very low traffic volumes, aging has a dominant effect and, as the traffic volume increases, the traffic and pavement strength have stronger effects, with the strongest traffic effects (steepest gradient) on the weakest pavements. In the case of surface treatments, cracking due to aging with very little traffic is indicated to occur after 13 years which, given a 67 percent probability of being within approximately ± 50 percent (see Table 4.7), is consistent with practical experience. The greater thickness of asphalt surfacings and the stiff base support of surfacings on cemented base pavements mean that those surfacings are less susceptible to 'punching' shear than thin surface treatments, and thus at low traffic volumes pavement strength or deflection provides an intercept effect which is not present for surface treatments.

Relationships D through H for maintenance surfacings are not fully statistically determined because of a limited data base. They relate the predicted time to the relationships for original surfacings (A to C), and to the extent of previous cracking which can appear by reflection in the new surfacing under certain conditions. For asphalt overlays on cemented base, the coefficients (KA and KW) of previous cracking assume reflection of cracking at propagation rates of 20 to 50 mm of layer thickness per year.

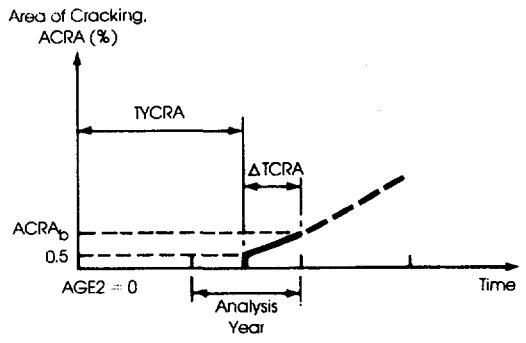
Cracking progression

The progression of cracked area was found in the empirical study to be a non-linear, S-shaped function, the rate of progression depending primarily on the area of cracking and the time since cracking initiation, without significant effects of either traffic loading or pavement strength (Paterson, 1987). The empirical evidence thus indicated that the rate of progression was a process related to the variability of the materials, with traffic and pavement strength influencing only the initiation of cracking. The function is symmetrical with the following general derivative form:

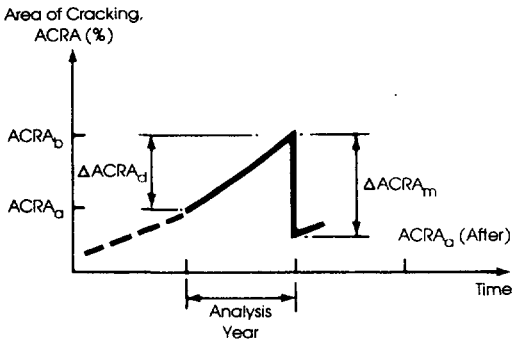
$$dA_t = a_i \, SA_t^{1-b_i} \, dt$$

Figure 4.4: Diagrammatic definition of primary variables in prediction of cracking initiation and progression

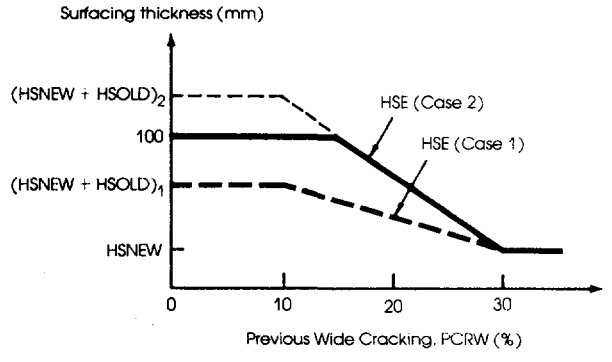
(a) Initiation variables



(b) Progression variables



(c) Effective surfacing thickness



Note: The definition of other surface distress variables is similar, substituting for ACRA and TYCRA as follows:
Wide cracking : ACRW, TYCRW Ravelling : ARAV, TYRAV

Source: This study.

Table 4.8: Models for predicting the initiation of all (i.e., narrow) cracking in various pavement types

Relationship	Pavement type
A	Surface treatments, granular base [1] $TYCRA = K_{ci} \{F_c\ RELIA + CRT\}$ where $RELIA = 13.2 \exp[-20.7(1 + CQ)\ YE4/SNC^2]$
B	All surfacings, cemented base (without stress-absorbing membrane) [1] $TYCRA = K_{ci} \{F_c\ RELIB + CRT\}$ where $RELIB = 1.12 \exp(.035\ HSE + .371\ \ell n\ CMOD - .418\ \ell n\ DEF - 2.87\ YE4\ DEF)$
DC	Asphalt concrete, granular base [1] $TYCRA = K_{ci} \{F_c\ RELIC + CRT\}$ where $RELIC = 4.21 \exp(0.14\ SNC - 17.1\ YE4/SNC^2)$
D	Slurry seal on surface treatment, granular base [2] $TYCRA = K_{ci} \{F_c\ RELID + CRT\}$ where $RELID = \max [RELIA \max (1 - PCRA/20, 0), 1.4]$
E	Reseal on surface treatment, granular base [2] $TYCRA = K_{ci} \{F_c\ RELIE + CRT\}$ where $RELIE = \max [RELIA \max (1 - PCRW/20, 0), 0.22\ HSNEW]$
F	Reseals on asphalt overlay, cemented base (without stress-absorbing membrane) [2] $TYCRA = K_{ci} \{F_c [(0.8\ KA + 0.2\ KW)(1 + 0.1\ HSE) + (1 - KA)(1 - KW)\ RELIB] + CRT\}$
G	Asphalt overlay on asphalt concrete, granular or bituminous base [2] $TYCRA = K_{ci} \{F_c\ RELIG + CRT\}$ where $RELIG = \max [RELIC \max (1 - PCRW/30,0), 0.025\ HSNEW]$
H	Surface treatment reseal on asphalt concrete, granular or bituminous base [2] $TYCRA = K_{ci} \{F_c\ RELIH + CRT\}$ where $RELIH = \max [RELIC \max (1 - PCRW/20,0), 0.12\ HSNEW]$

[1] Statistically derived from Brazil-UNDP road deterioration study.
[2] Empirically developed based on Brazil-UNDP study data and judgment.
Source: Adapted from Paterson (1987).

Table 4.9: Models for predicting the initiation of wide cracking in various pavement types

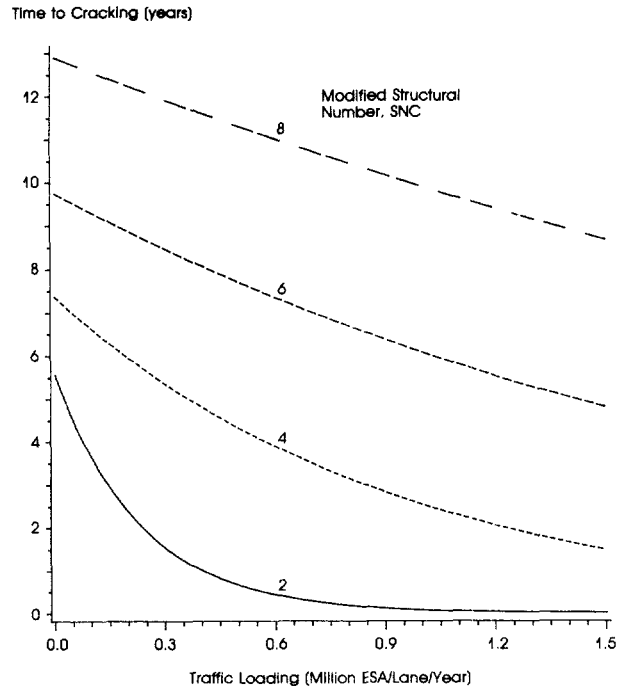
Relationship	Pavement type and model
A	Surface treatments, granular base [1]
	$TYCRW = K_{ci} \max(2.66 + 0.88\ TYCRA,\ 1.16\ TYCRA)$
B	All surfacings, cemented base (without stress-absorbing membrane) [1]
	$TYCRW = K_{ci}\ (1.46 + 0.98\ TYCRA)$
C	Asphalt concrete, granular base [1]
	$TYCRW = K_{ci}\ (2.46 + 0.93\ TYCRA)$
D	Slurry seal on surface treatment [1]
	$TYCRW = K_{ci}\ (0.70 + 1.65\ TYCRA)$
E, H	All surface treatment reseals, granular[1] or bituminous base [2]
	$TYCRW = K_{ci}\ (1.85 + TYCRA)$
F	Reseals or asphalt overlay on cemented base (without stress-absorbing membrane [1]
	$TYCRW = K_{ci}\ 1.78\ TYCRA$
G	Asphalt overlay on asphalt concrete, granular[1] or bituminous base [2]
	$TYCRW = K_{ci}\ (2.04 + 0.98\ TYCRA)$

[1] Statistically derived from Brazil-UNDP road deterioration study.
[2] Empirically developed based on Brazil-UNDP study data and judgment.
Source: Paterson (1987).

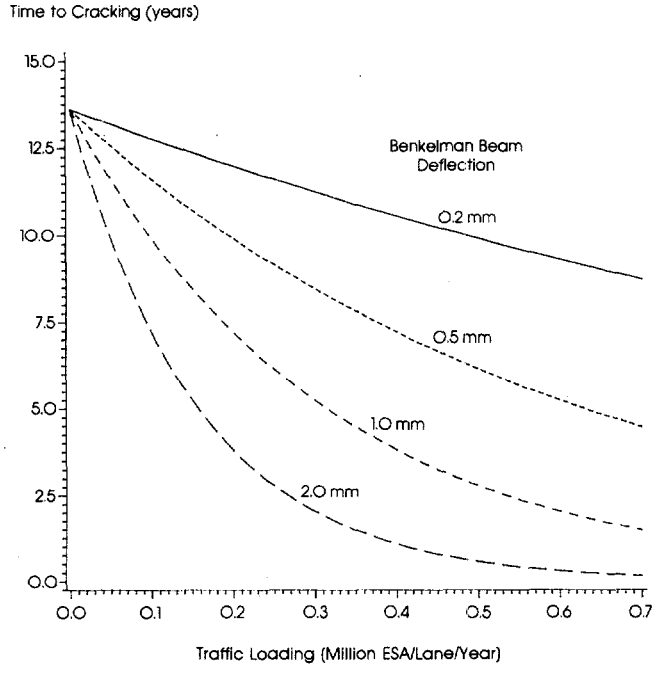
where A_t is the area distressed at time t; a_i, b_i are coefficients estimated empirically for pavement type i (see Table 4.10) and SA_t is a function of A_t symmetrical about an area of 50 percent as defined below for each relationship. The functions for each pavement type are illustrated in Figure 4. In the model the relationships have a general first-difference form as follows:

Figure 4.5: Predictions of time cracking initiation for original pavements with granular base

(a) Asphalt concrete surfacing



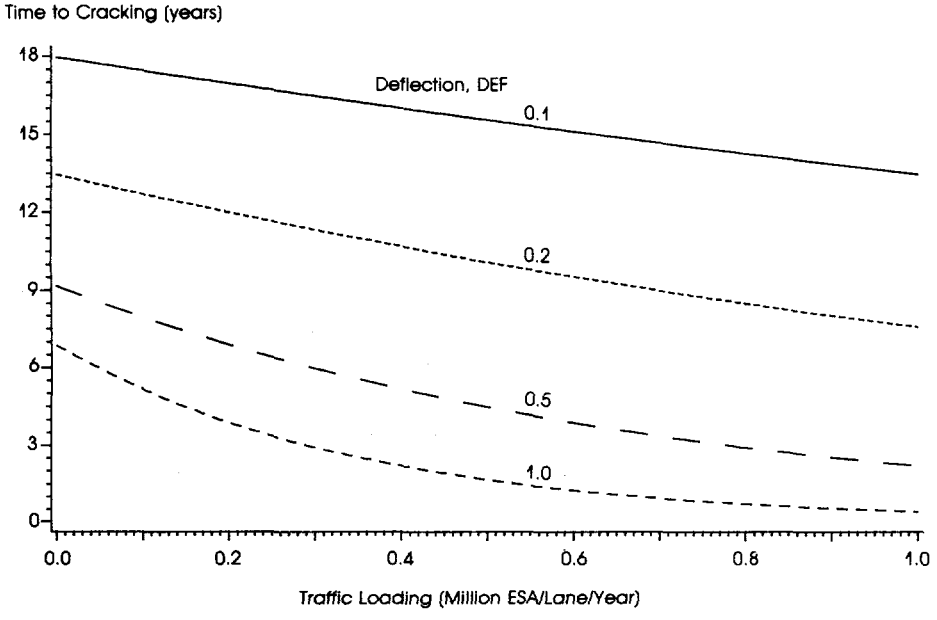
(b) Double surface treatment surfacings



Source: Paterson (1987).

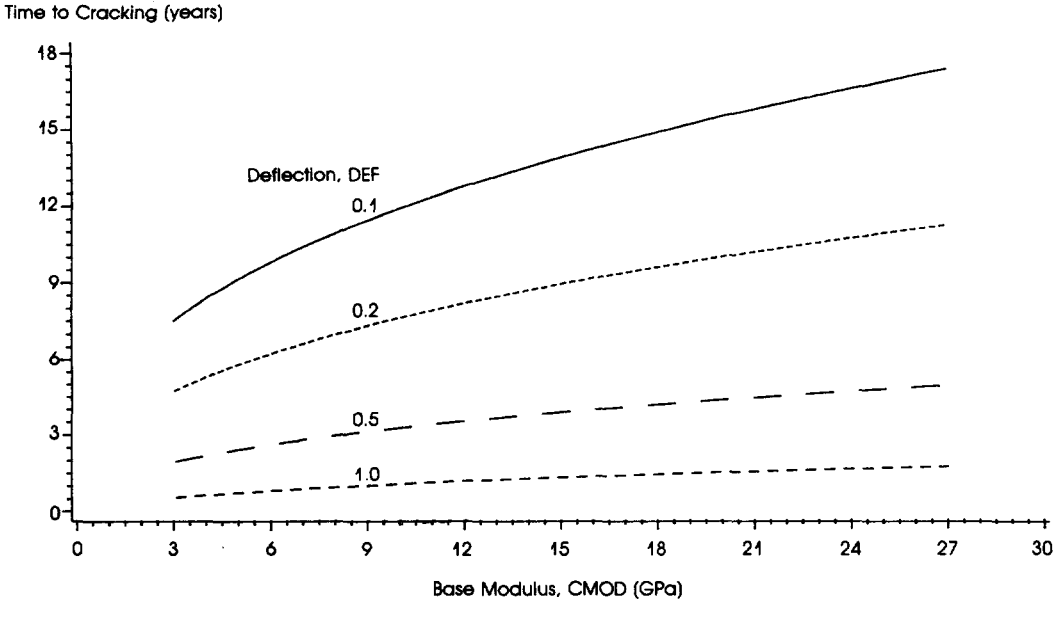
Figure 4.6: Predictions of time cracking initiation for original pavements with cemented base

(a) Related to traffic loading



Note: Modulus of base 20 GPa; Surfacing thickness 20 mm.

(b) Related to base stiffness



Note: Traffic 0.5 million ESA/lane/yr; Surfacing thickness 20 mm.
Source: Paterson (1987).

1. All-cracking area:

$$\Delta ACRA_d = K_{cp}\ CRP\ z_a\ \{[z_a\ a_i\ b_i\ \Delta TCRA + SCRA_a^{b_i}]^{1/b_i} - SCRA_a\}$$

2. Wide-cracking area:

$$\Delta ACRW_d = K_{cp}\ CRP\ z_w\ \{[z_w\ c_i\ d_i\ \Delta TCRW + SCRW_a^{d_i}]^{1/d_i} - SCRW_a\}$$

where $\Delta ACRA_d$ = the predicted change in the area of all cracking during the analysis year due to road deterioration, in percent of the total carriageway area;

$\Delta TCRA$ = the fraction of the analysis year in which all-cracking progression applies, in years, given by:

$$\Delta TCRA = \begin{cases} 0 & \text{if } AGE2 < TYCRA \text{ and } ACRA_a = 0 \\ (AGE2 - TYCRA) & \text{if } AGE2 - 1 < TYCRA \leq AGE2 \text{ and } ACRA_a = 0 \\ 1 & \text{if } TYCRA \leq AGE2 - 1 \text{ or } ACRA_a > 0 \end{cases}$$

$$SCRA_a = \min(ACRA_a,\ 100 - ACRA_a)$$

$$SCRA_a = \max(SCRA_a,\ 0.5) \text{ if } ACRA_a > 0.5$$

$$ACRA_a = \begin{cases} 0 & \text{if } \Delta TCRA = 0 \\ 0.5 & \text{if } 0 < \Delta TCRA < 1 \\ ACRA_a & \text{otherwise} \end{cases}$$

$\Delta ACRW_d$ = the predicted change in the area of wide cracking during the analysis year due to road deterioration, in percent of the total carriageway area.

$\Delta TCRW$ = fraction of the analysis year in which wide cracking progression applies, in years, given by:

$$\Delta TCRW = \begin{cases} 0 & \text{if } AGE2 < TYCRW \text{ and } ACRW_a = 0 \\ (AGE2 - TYCRW) & \text{if } AGE2 - 1 < TYCRW \leq AGE2 \text{ and } ACRW_a = 0 \\ 1 & \text{if } TYCRW \leq AGE2 - 1 \text{ or } ACRW_a > 0 \end{cases}$$

$$SCRW_a = \min(ACRW_a,\ 100 - ACRW_a)$$

$$SCRW_a = \max(SCRW_a,\ 0.5) \text{ if } ACRW_a > 0.5$$

$$ACRW_a = \begin{cases} 0 & \text{if } \Delta TCRW = 0 \\ 0.5 & \text{if } 0 < \Delta TCRW < 1 \text{ and } ACRW <= 0.5 \\ ACRW_a & \text{otherwise} \end{cases}$$

Table 4.10: Coefficients for prediction of cracking progression for all pavement types

Relationship	Pavement type [1]	a	b	c	d
C	Asphalt concrete	1.84	0.45	2.94	0.56
A	Surface treatment	1.76	0.32	2.50	0.25
B, F	ST on cemented base	2.13	0.35	3.67	0.38
G	Asphalt overlays	1.07	0.28	2.58	0.45
E, H	Reseals	2.41	0.34	3.4	0.35
D	Slurry seals	2.41	0.34	3.4	0.35

[1] Surfacing as indicated on flexible (granular or bituminous) base unless otherwise stated.
Source: Paterson (1987).

a_i, b_i, c_i, d_i = coefficients for pavement type i given in Table 4.10

K_{cp} = the user-specified deterioration factor for cracking progression (default value = 1); and

CRP = the retardation of cracking progression due to preventive treatment, given by CRP = 1 - 0.12 CRT

z_a = 1 if $ACRA_a < 50$; $z_a = -1$ otherwise;

z_w = 1 if $ACRW_a < 50$; $z_w = -1$ otherwise.

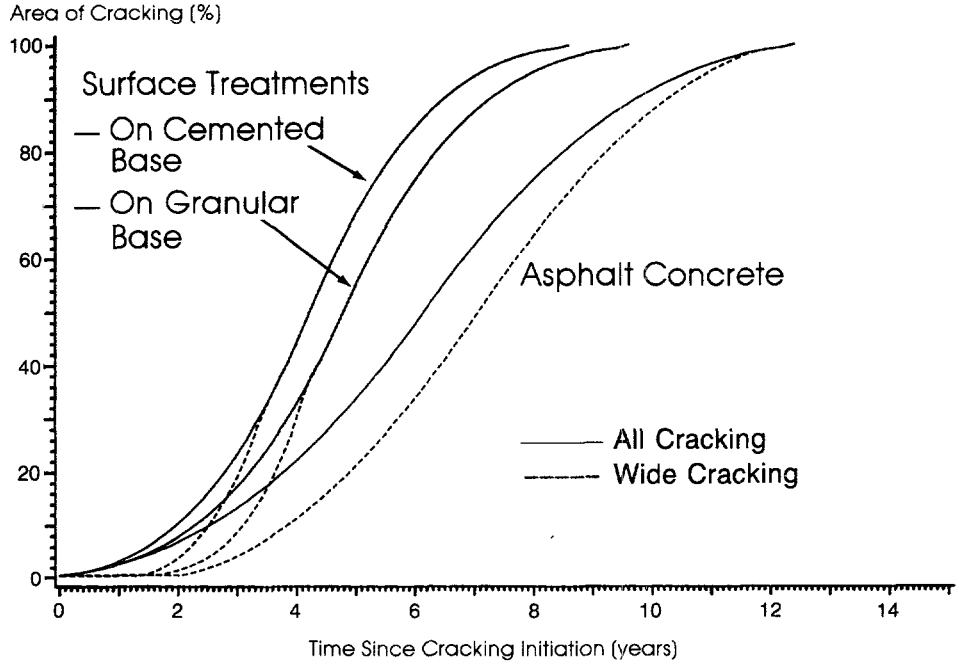
In addition to the above basic definitions, the initiation of wide cracking is constrained so that it does not commence before the area of all cracking ($ACRA_a$) exceeds 5 percent. In particular, this prevents wide and narrow cracking commencing simultaneously after patching, when both ΔTCRA and ΔTCRW equal 1. Furthermore, after the patching of wide cracking, the rate of progression of wide cracking is assumed to begin not at the low initial rate given by $ACRW_a = 0.5$ but at a higher rate equivalent to its rate before patching, here simulated by defining a temporary value of wide cracking, X to be 5 percent less than ACRA, i.e.:

$\Delta TCRW = 0$ if $ACRA_a <= 5$ and $ACRW_a <= 0.5$ and $\Delta TCRW > 0$.

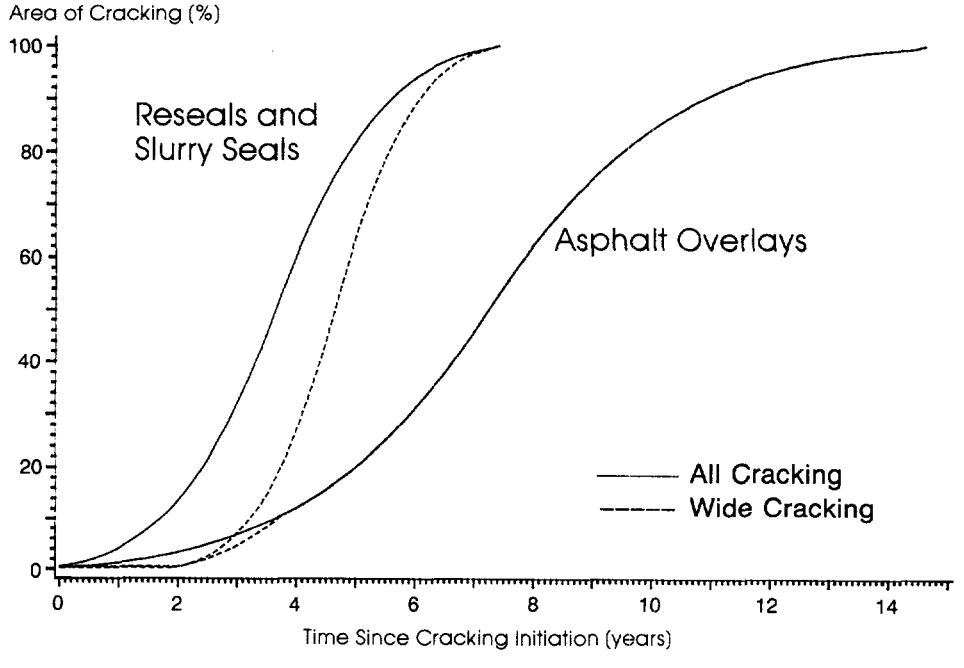
$X = ACRA_a - 5$ if $ACRW_a <= 1$ and $ACRA_a > 11$.

Figure 4.7: Predictions of cracking progression for original, reseal, and overlay surfacings

(a) Original surfacings



(b) Reseal and overlay surfacings



Source: Paterson (1987).

where X is a temporary value of $ACRW_a$ used only in the computation of $\Delta ACRW_d$ above.

4.2.3 Ravelling Initiation and Progression

Ravelling is the loss of surfacing material from pavements, which in thin surfacings may eventually develop into potholes. Usually it is caused by weathering of the binder, and except in very old pavements is usually limited to the upper layer of double surface treatments.

The empirical models for predicting the initiation and progression of ravelling derived from the Brazil study relate to both original and maintenance surfacings of double surface treatments, slurry seals and open-graded cold-mix asphalt, but not to asphalt concrete. The relationships are defined in Table 4.11 and illustrated in Figure 4.8. They have probabilistic and incremental forms similar to the cracking models. Structural properties are not significant, and the only explanatory variables are the surfacing type and construction quality. Major construction faults such as faulty binder distribution, contaminated stone, stripping of the binder, etc., cause a 50 percent reduction in "life" before ravelling. In progression, no variables except time and area of ravelling were significant in the empirical study. The supplementary variables are:

TYRAV = the predicted number of years to ravelling initiation since the last surfacing or resurfacing (when the surfacing age AGE2 = 0);

K_{vi} = the user-specified deterioration factor for ravelling initiation (default value = 1);

F_r = the occurrence distribution factor for ravelling initiation for the subsection (the values used in HDM-III are 0.54, 0.97 and 1.49 for the weak, medium and strong subsections, respectively);

$\Delta ARAV_d$ = the predicted change in the ravelled area during the analysis year due to road deterioration, in percent; and

ΔTRAV = the fraction of the analysis year during which ravelling progression applies, in years, given by:

$$\Delta TRAV = \begin{cases} 0 & \text{if } AGE2 < TYRAV \text{ and } ARAV_a = 0 \\ (AGE2 - TYRAV) & \text{if } AGE2 - 1 < TYRAV \leq AGE2 \text{ and } ARAV_a = 0 \\ 1 & \text{if } TYRAV \leq AGE2 - 1 \text{ or } ARAV_a > 0 \end{cases}$$

$$SRAV_a = \min(ARAV_a, 100 - ARAV_a)$$

$$ARAV_a = \begin{cases} 0.5 & \text{if } 0 < \Delta TRAV < 1 \text{ and } ARAV_a <= 0.5 \\ ARAV_a & \text{otherwise} \end{cases}$$

Table 4.11: Models for predicting the initiation and progression of ravelling of various surfacings

Relationship	Pavement type and model

RAVELLING INITIATION

A Surface treatments including reseals (ST, RSST, RSAC) [1]

$$TYRAV = K_{vi} \{F_r [10.5 \exp(-0.655\ CQ - 0.156\ YAX)]\ RRF\}$$

B Slurry seal on surface treatment or asphalt concrete (SSST) [1]

$$TYRAV = K_{vi} \{F_r [14.1 \exp(-0.655\ CQ - 0.156\ YAX)]\ RRF\}$$

C Cold-mix surfacing or cold-mix overlay (OCMS) [1]

$$TYRAV = K_{vi} \{F_r [8.0 \exp(-0.655\ CQ - 0.156\ YAX)]\ RRF\}$$

D Asphalt concrete and asphalt overlays (AC, OVSA) [2]

$$TYRAV = 100$$

RAVELLING PROGRESSION

All surface treatments, reseals, slurry seal, cold-mix (ST, RSST, RSAC, SSST, OCMS [1]

$$\Delta ARAV_d = (K_{vi}\ RRF)^{-1}\ z_r \{[z_r\ 1.560\ \Delta TRAV + SRAV_a^{0.352}]^{2.84} - SRAV_a\}$$

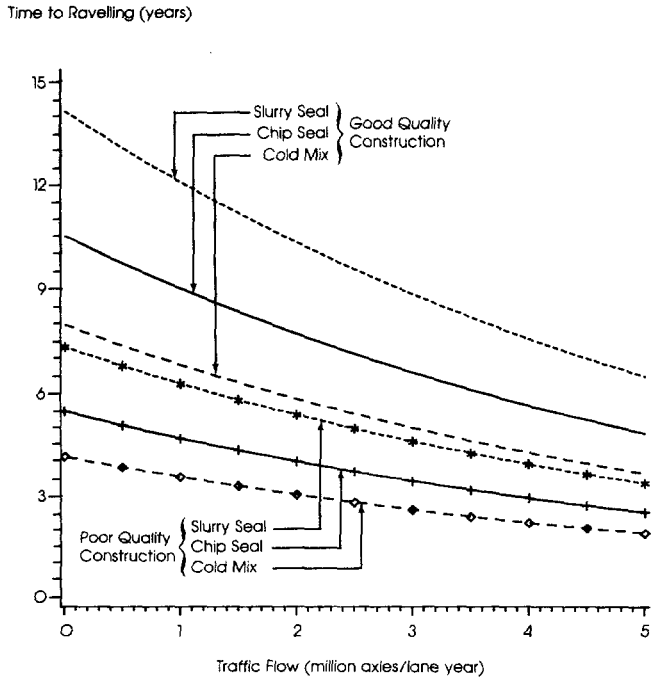
Asphalt concrete and asphalt overlays (AC, OVSA) [2]

$$\Delta ARAV_d = 0$$

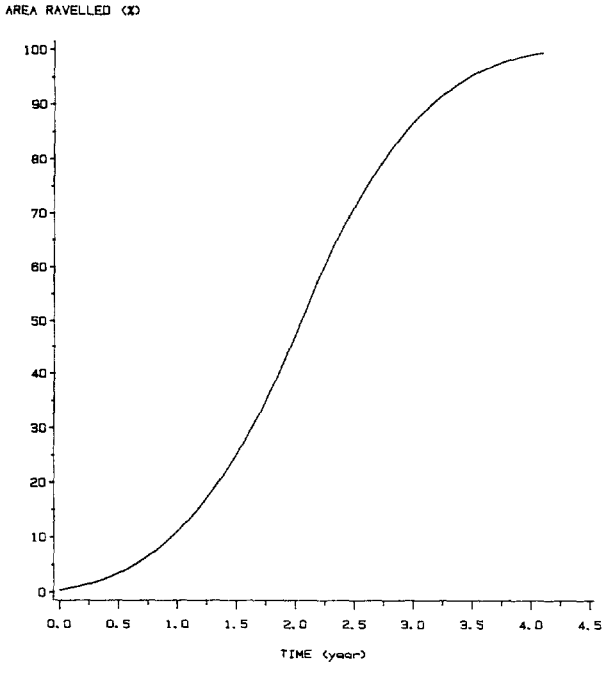
[1] Statistically derived from Brazil-UNDP road deterioration study.
[2] Default relationship assuming sound specification and construction of asphalt mixture.
Source: After Paterson (1987).

Figure 4.8: Predictions of ravelling initiation and progression for various thin surfacings on flexible and semi-rigid pavements

(a) Ravelling initiation



(b) Ravelling progression



Source: After Paterson (1987).

$z_r = 1$ if $ARAV_a < 50$; $z_r = -1$ otherwise.

4.2.4 Potholing Initiation and Progression

Potholing usually develops from the spalling of wide cracking or the ravelling of thin surface treatments, although in new surface treatment construction it can develop from random local defects in the surfacing or base. In both cases the initiation and progression of potholes are highly dependent on the mechanical disintegration properties of the base. As these are impractical to quantify for a network analysis, the base type and modified structural number are used as surrogate measures. The relationships are empirical, based on data from the Caribbean, Ghana, Brazil and Kenya studies, because reliable statistical estimation was impracticable (Paterson, 1987).

Initiation is expressed as a function of the time (TMIN) since the initiation of triggering distress and traffic flow, and occurs typically 2 to 6 years after wide cracking and 3 to 6 years after ravelling of thin surface treatments, as shown in Figure 4.8 (a) and given by:

$$TMIN = \begin{cases} \max\,[6 - YAX;\ 2] & \text{if base is cemented} \\ \max\,[2 + 0.04\ HS - 0.5\ YAX;\ 2] & \text{otherwise} \end{cases}$$

$$HS = \begin{cases} (HSNEW + HSOLD) & \text{if base is granular} \\ (HSNEW + HSOLD + HBASE) & \text{if base is bituminous.} \end{cases}$$

The initiation indicator INPOT is set to 1 either if the age of the oldest relevant distress remaining after maintenance estimated from the rate of progression is greater than TMIN, or if potholing is present in the current condition or if potholing was present in the first analysis year and there has been no subsequent reseal, overlay reconstruction or construction, as follows:

$$INPOT = \begin{cases} 1 & \text{if } AGE2\text{-}TYCRW \geq TMIN \text{ and } ACRW_d \geq 20 \\ & \text{or } AGE2\text{-}TYRAV \geq TMIN + (HSNEW + HSOLD)/10 \\ & \quad \text{and } ARAV_d \geq 30 \\ & \text{or } APOT_a > 0 \\ & \text{or } APOT_a(\text{first analysis year}) > 0 \\ & \quad \text{and } [AGE1 - AGE1\ (\text{first analysis year}) = \\ & \quad (\text{analysis year} - \text{first analysis year}) \\ 0 & \text{otherwise.} \end{cases}$$

The progression of potholing is computed firstly as a volume, in cubic meters per lane-km, because the effect on roughness has been shown to be linearly related to pothole volume (Paterson, 1987). For consistency with the accounting of other distress types, potholing is then converted to an equivalent percentage area, assuming an average pothole depth of 80 mm, by the factor (0.8 W/ELANES) m^3/lane-km per percent area. The potholing progression comprises three components, i.e., new potholes caused by wide cracking ($\Delta APOTCR_d$), new potholes caused by ravelling ($\Delta APOTRV_d$) and the enlargement of existing potholes ($\Delta APOTP_d$), as illustrated in Figure 4.9 (b) and (c) and defined as follows:

$$\Delta APOT_d = \min \{\Delta APOTCR_d + \Delta APOTRV_d + \Delta APOTP_d;\ 10\}$$

where $\Delta APOT_d$ = the predicted change in the total area of potholes during the analysis year due to road deterioration, in percent;

$\Delta APOTCR_d$ = the predicted change in the area of potholes during the analysis year due to cracking, given by:

$$\Delta APOTCR_d = \begin{cases} K_{pp}\ INPOT \min [2ACRW_a\ U;\ 6] \\ \quad \text{if } ACRW_a > 20 \\ 0 \text{ otherwise} \end{cases}$$

$$U = (1+CR)\ (YAX/SNC)/((HSNEW + NSOLD)\ 0.8W/ELANES);$$

$\Delta APOTRV_d$ = the predicted change in the area of potholes during the analysis year due to ravelling, given by:

$$\Delta APOTRV_d = \begin{cases} K_{pp}\ INPOT \min [0.4\ ARAV_a\ U;\ 6] \\ \quad \text{if } ARAV_a > 30 \\ 0 \text{ otherwise} \end{cases}$$

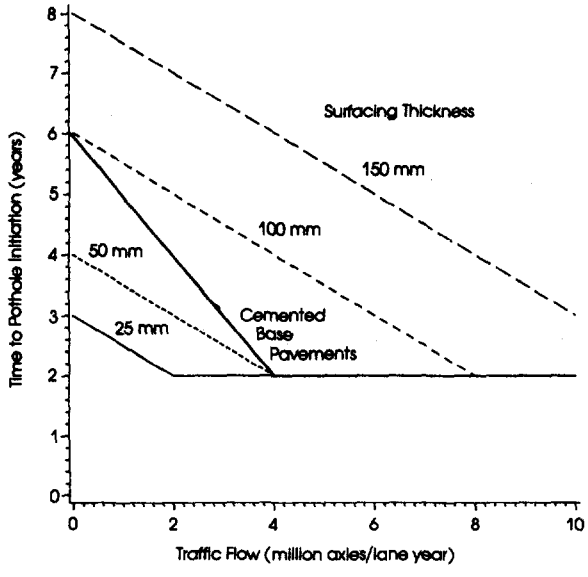
$\Delta APOTP_d$ = the predicted change in the equivalent area of potholes during the analysis year due to enlargement, given by:

$$\Delta APOTP_d = \min \{APOT_a\ [KBASE\ YAX\ (MMP + 0.1)];\ 10\}$$

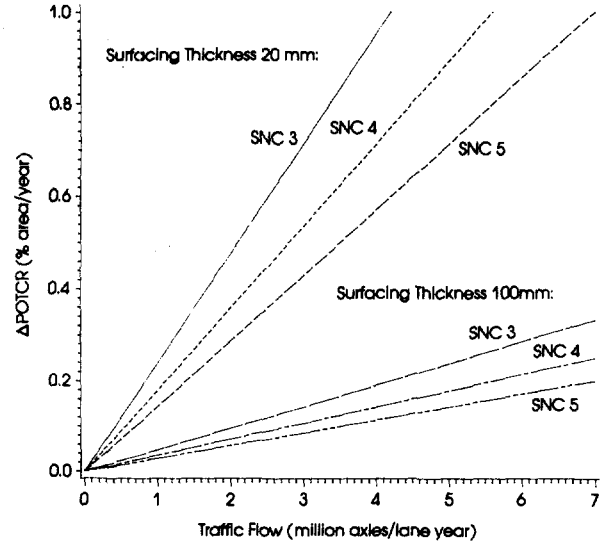
$$KBASE = \begin{cases} \max [2 - 0.02\ (HSNEW + HSOLD);\ 0.3] & \text{if base = granular} \\ 0.6 & \text{if base = cement-treated} \\ 0.3 & \text{otherwise} \end{cases}$$

Figure 4.9: Parameter variations for prediction of potholing initiation and progression

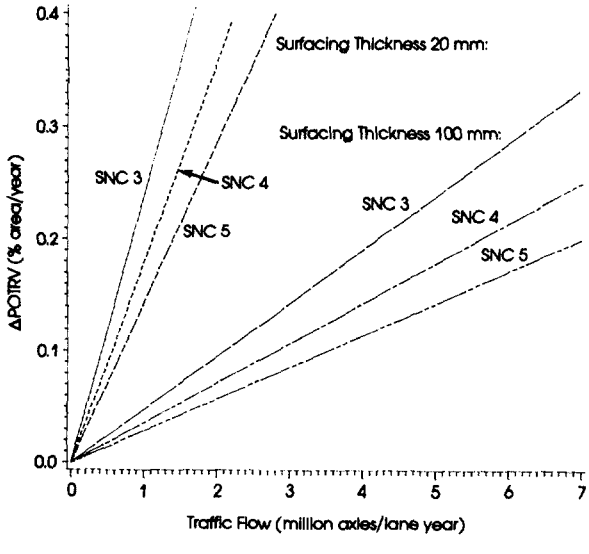
(a) Initiation



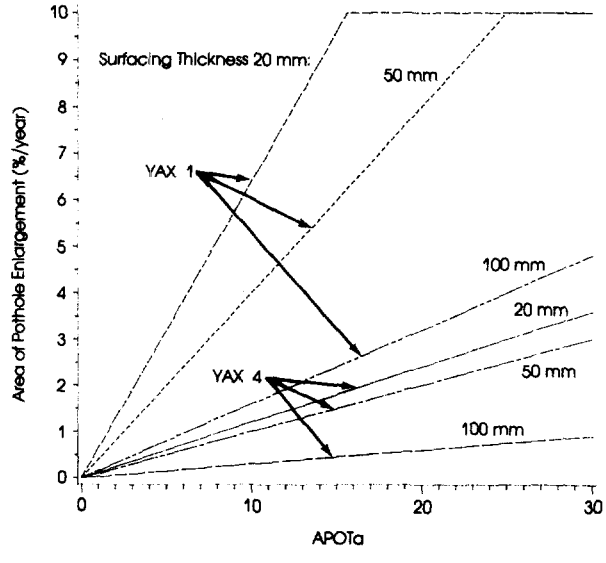
(b) Annual increment due to cracking



(c) Annual increment due to ravelling



(d) Enlargement



and K_{pp} = the user-specified deterioration factor for pothole progression (default value = 1) (Section 4.1.8).

4.2.5 Surface Damage at the End of the Year before Maintenance

Using the quantities computed above and appropriate limits, the damaged areas at the end of the analysis year before maintenance are predicted in the manner described below.

For modelling purposes the road surface is assumed to consist of the areas potholed, cracked, ravelled and undamaged. The undamaged area consists of the original road area which is still in intact condition since the last surfacing or resurfacing and the area which has been patched. These areas are assumed to be mutually exclusive, so the sum of the potholed, cracked, ravelled and undamaged areas must equal 100 percent. In devising a logic that satisfies this constraint the following simplifying assumptions are made:

1. Cracking develops first from the undamaged area and then, after the latter is exhausted, from the ravelled area. Furthermore, an area once cracked can develop potholes but cannot ravel.

2. Ravelling can only develop from the undamaged area, and after an area is ravelled it also can crack, at which stage it is reclassified from ravelled to cracked.

3. Potholes can only develop from cracked, ravelled and undamaged areas (as reflected in the formulas for computing the change in potholed area), and unless it is repaired, an area potholed cannot revert to cracking, ravelling or undamaged status.

4. An upper limit of 30 percent is imposed on the potholed area because above this level the pavement surface becomes ill-defined and the roughness function becomes invalid.

The above assumptions lead to the following equations for computing the damaged areas before maintenance:

$$APOT_b = \min(30, APOT_a + \Delta APOT_d)$$
$$ACRA_b = \min(100 - APOT_b; ACRA_a + \Delta ACRA_d - \Delta APOTCR_d)$$
$$ACRW_b = \min(100 - APOT_b; ACRW_a + \Delta ACRW_d - \Delta APOTCR_d; ACRA_b)$$
$$ARAV_b = \min(100 - APOT_b - ACRA_b; ARAV_a + \Delta ARAV_d - \Delta APOTRV_d)$$

4.2.6 Rut Depth Progression

All surface and base types employ the same relationships for predicting rut depth progression. Mean rut depth is not used as a maintenance intervention criterion in HDM-III, but is used as a means to estimate the variation of rut depth (standard deviation) which contributes directly to roughness.

Mean rut depth

The progression of mean rut depth is predicted by the following relationship (Paterson, 1985):

$$RDM = K_{rp} \frac{39800 \left[YE4\ 10^6\right]^{ERM}}{SNC^{0.502}\ COMP^{2.30}} .$$

In the first year when $RDM_a = 0$ the equation above is used directly to estimate ΔRDM_d), but subsequently the incremental change in mean rut depth due to road deterioration during the analysis year is derived from this as follows.

$$\Delta RDM_d = K_{rp} \left[\frac{0.166 + ERM}{AGE3} + 0.0219\ MMP\ \Delta CRX_d\ \ell n\ [\max (1;\ AGE3\ YE4)] \right] RDM_a \quad \text{if } RDM_a > 0$$

where ΔRDM_d = the predicted change in the mean rut depth during the analysis year due to road deterioration, in mm;

K_{rp} = the user-specified deterioration factor for rut depth progression (default value = 1);

ERM = the exponent which is a function of surface characteristics and precipitation, given by:

$$ERM = 0.09 - 0.0009\ RH + 0.0384\ DEF + 0.00158\ MMP\ CRX_a;$$

CRX_a = area of indexed cracking at the beginning of the analysis year, given by:

$$CRX_a = 0.62\ ACRA_a + 0.39\ ACRW_a;$$

and ΔCRX_d = the predicted change area of indexed cracking due to road deterioration, given by:

$$\Delta CRX_d = 0.62\ (ACRA_b - ACRA_a) + 0.39\ (ACRW_b - ACRW_a)$$

The mean rut depth at the end of the year, with a limit of 50 mm, is given by:

$$RDM_b = \min\ (50;\ RDM_a + \Delta RDM_d).$$

The prediction of mean rut depth is illustrated in Fig. 4.10 which shows a diminishing rate of increase with time and a slight effect of rainfall and cracking. The level of compaction has a strong influence and the rapid rise in rut depth initially is probably due to this influence. Unless the structural number is very low for the traffic to be carried, rutting is not usually a major problem with modern pavement construction specifications. The problem of rutting which may develop in thick or soft asphalt layers in high temperatures is not covered by the prediction model above.

Standard deviation of rut depth

The variation of rut depth which directly influences roughness is expressed by the standard deviation and is a strong function of the mean rut depth as follows (Paterson, 1987):

$$RDS = \frac{K_{rp}\ 4390\ RDM_d^{0.532}\ (YE4\ 10^6)^{ERS}}{SNC^{0.422}\ COMP^{1.66}}\ .$$

In the first year when $RDM_a = 0$, the equation above is used directly to estimate ΔRDS_d where $RDM = \Delta RDM_d$, and the prediction is halved as an explicit suppression of the sharp initial increase. Subsequently the incremental change in rut depth standard deviation due to road deterioration in the analysis year is derived from the above equation as follows

$$\Delta RDS_d = K_{rp} \left[\frac{0.532\ (RDM_b - RDM_a)}{RDM_a} + \frac{ERS}{AGE3} + 0.0159\ MMP\ \Delta CRX_d \right.$$

$$\left. \ell n\ [\max\ (1;\ AGE3\ YE4)] \right] RDS_a \qquad \text{if } RDM_a > 0 \text{ and } RDS_a > 0$$

where ΔRDS_d = the predicted change in the standard deviation of rut depth during the analysis year due to road deterioration, in mm; and

ERS = an exponent which is a function of the surface characteristics and precipitation, given by:

$$ERS = -0.0086\ RH + 0.00115\ MMP\ CRX_a.$$

The standard deviation of rut depth at the end of the year, with an upper limit equal to the mean rut depth, is given by:

$$RDS_b = \min\ (RDM_b;\ RDS_a + \Delta RDS_d)$$

The strong relationship between the standard deviation and mean of rut depth is illustrated in Fig. 4.10(b).

4.2.7 Roughness Progression

Roughness progression is predicted as the sum of three components: structural deformation related to roughness, equivalent standard axle load flow, and structural number; surface condition, related to changes in cracking, potholing and rut depth variation; and an age-environment-related roughness term. All flexible pavement types employ the same relationship, estimated from the Brazil study, for predicting the incremental change in roughness due to deterioration but excluding maintenance effects (see Section 4.3 later), given by (Paterson, 1987):

$$\Delta QI_d = 13\, K_{gp}\Big[134\ EMT\ (SNCK + 1)^{-5.0}\ YE4 + 0.114\ (RDS_b - RDS_a) + 0.0066\ \Delta CRX_d + 0.42\ \Delta APOT_d\Big] + K_{ge}\ 0.023\ QI_a$$

where ΔQI_d = the predicted change in road roughness during the analysis year due to road deterioration, in QI;

K_{gp} = the user-specified deterioration factor for roughness progression (default value = 1);

K_{ge} = the user-specified deterioration factor for the environment-related annual fractional increase in roughness (default value = 1);

EMT = exp (0.023 K_{ge} AGE3)

SNCK = the modified structural number adjusted for the effect of cracking, given by:

$$SNCK = \max\ (1.5;\ SNC - \Delta SNK)$$

ΔSNK = the predicted reduction in the structural number due to cracking since the last pavement reseal, overlay or reconstruction (when the surfacing age, AGE2, equals zero), given by:

$$\Delta SNK = 0.0000758\ [CRX'_a\ HSNEW + ECR\ HSOLD]$$

CRX'_a = min (63; CRX_a);

ECR = the predicted excess cracking beyond the amount that existed in the old surfacing layers at the time of the last pavement reseal, overlay or reconstruction, given by:

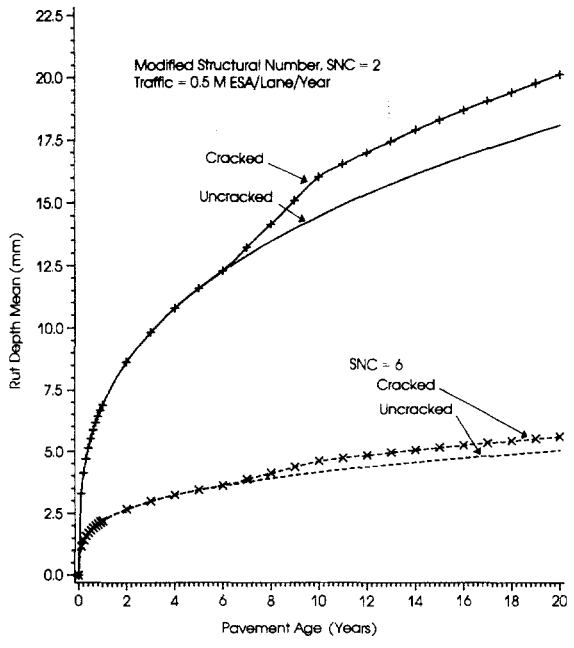
$$ECR = \max\ [\min(CRX_a - PCRX;\ 40);\ 0]$$

PCRX = area of previous indexed cracking in the old surfacing and base layers, given by:

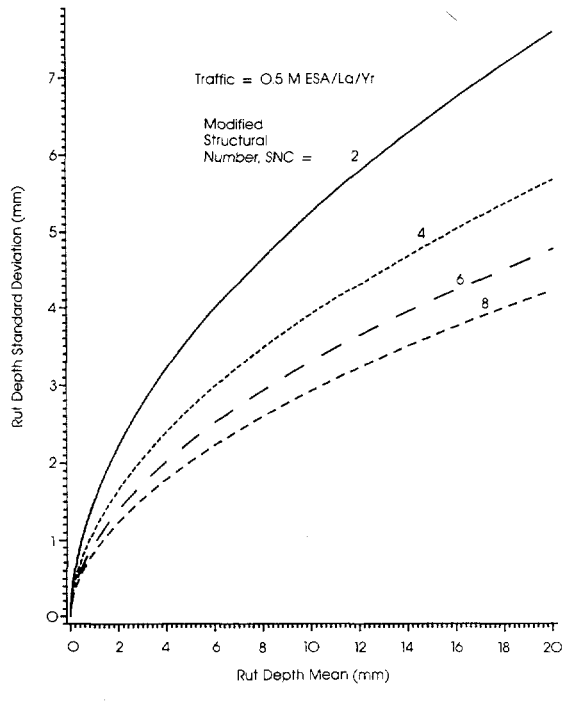
$$PCRX = 0.62\ PCRA + 0.39\ PCRW.$$

Figure 4.10: Predictions of the progression of the mean and standard deviation of rut depth

(a) Mean rut depth



(b) Standard deviation of rut depth



Source: Paterson (1987)

Roughness at the end of the analysis year, before maintenance and imposing an upper limit of 150 QI, is given by:

$$QI_b = \min (150; QI_a + \Delta QI_d).$$

Predictions from the model are illustrated in Fig. 4.11 for two pavements (SNC-values of 3 and 5) under six volumes of traffic loading and minimal maintenance comprising the patching of all potholes. At the lowest traffic loadings, the roughness progression is a function primarily of the last term in the model representing age-environment effects, which amounted to 2.3 percent/year for the warm, subhumid to humid climate of the Brazil empirical base and may need adjustment for other climate and soil-type regions through the factor, K_{ge}. Traffic loading and pavement strength effects are given by the explicit traffic-structural number-age term and also through the rut depth variation and cracking terms; adjustment of these coefficients can be made uniformly through the factor K_{gp} although this is considered unlikely to be necessary.

4.3 PAVED ROAD MAINTENANCE INTERVENTION

4.3.1 Classification and Hierarchy

Maintenance operations are classified into six primary categories for the submodel, based on when they are to be applied and their impact on pavement condition and strength. These are listed in Table 4.12 together with a summary of the options available for intervention criteria, the types of maintenance within each category, and the hierarchy with which they are applied in the model. Detailed description is given in the following sections. The intervention criteria determine the timing and type of maintenance to be applied and together comprise the maintenance standard specified by the user for each maintenance category. The timing of intervention is specified for each type of maintenance individually as either:

1. Scheduled, that is a fixed amount per year (e.g., m^2/km) or at fixed intervals of time (e.g., every 3 years); or

2. Condition-responsive, that is intervening when the pavement condition is predicted to reach a critical threshold level which is specified by the user.

These two types represent the majority of maintenance policies, either explicitly or implicitly, although in practice there is often a blend of the two. Hence the condition-responsive option has been extended to include user-specified limits on the minimum and maximum intervals between successive treatments. For example, it is possible to specify resealing to be applied when the total damaged area reaches 30 percent, but to limit the timing so that the resealing would not be done any earlier than, say, 4 years after the previous major maintenance (perhaps to minimize traffic disruptions), and no later than, say, 15 years (a possible preventive policy).

In order to simulate realistic policies further, it is recognized that practical considerations would normally preclude the situation of periodic maintenance being applied immediately before a major overlay or reconstruction were planned. Provision is therefore made to specify:

1. The latest time; and

2. The maximum applicable roughness at which each type of maintenance might be performed.

This also dictates a natural hierarchy in which lesser treatments cannot be performed in the same year as a major treatment.

Thus, in the first year after construction or reconstruction, no maintenance operations are performed except the routine-miscellaneous operation. Routine-miscellaneous operations are performed every year regardless of other operations. In any analysis year only one of the five operations, 1 to 5 in Table 4.12, can be performed in addition to the routine-miscellaneous operation. An exception to that rule is that the resealing operation includes an option for preparatory patching. Otherwise, preparatory work such as preliminary patching, crack sealing or a levelling course, are expected to be included in the unit costs of the operation.

The hierarchical order simulated in the maintenance decision is from 1 to 5 in the table, excluding routine-miscellaneous. Pavement reconstruction is performed in an analysis year if its criteria are satisfied irrespective of the criteria for the remaining operations. Otherwise, the criteria for overlaying are considered next, and if satisfied, overlaying is performed irrespective of the criteria for resealing, preventive treatment, patching, etc. And so the hierarchy continues.

Except for routine-miscellaneous and patching operations, which are assumed to be performed regularly during the year, all periodic, major maintenance is assumed for simplicity to be performed at the end of the analysis year.

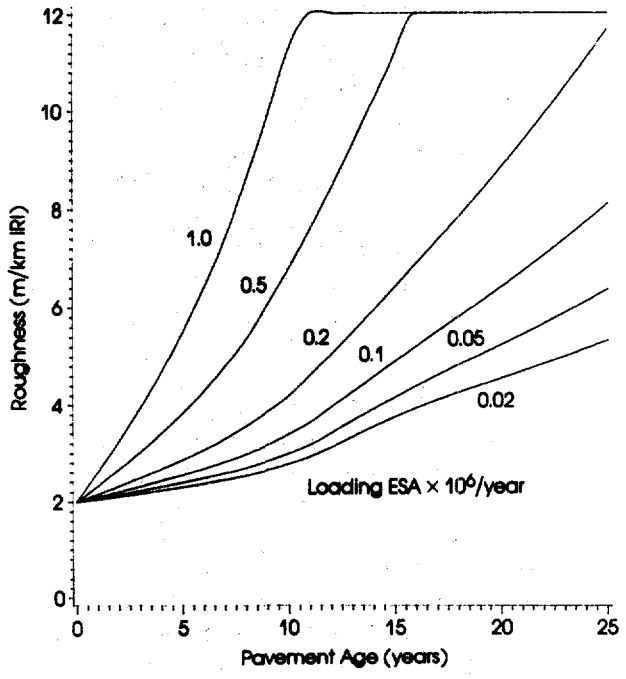
It should be noted that, in this version, all maintenance operations are specified explicitly by the user and that no endogeneous design or selection from alternative options is made by the submodel. For example, the user must specify the thickness of overlay or reseal to be applied. The effect of the maintenance on pavement condition, strength and history, however, is computed endogeneously. Patching areas for example are calculated by applying the user's specified maintenance policy to the damaged areas predicted by the model.

4.3.2 Routine-Miscellaneous Maintenance

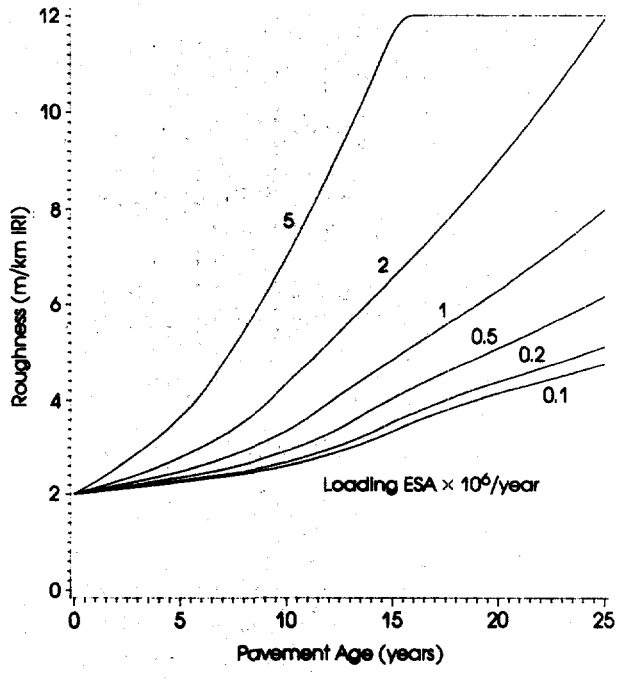
The HDM model does not model the effects of alternative levels of routine maintenance other than pothole patching. Rather for vital components, i.e., drainage features and shoulders, the model assumes levels

Figure 4.11: Prediction of roughness progression for flexible and semi-rigid pavements under minimal maintenance of patching all potholes

(a) Asphalt Concrete Modified Structural Number 3



(b) Asphalt Concrete Modified Structural Number 5



Source: Paterson (1987).

Table 4.12: Classification and hierarchy of maintenance for paved roads

Class	Hierarchy	Frequency	Intervention Criteria (IC)	Type options	Effects
Routine-miscellaneous	6	Annual, automatic	Nil	Nil	Nil; absence would indicate negative effects.
Patching	5	Periodic, specified by IC, % area, fixed area, or limit	1. Scheduled if damaged, limited area 2. Condition a) all surface b) potholes only	a. Surface patching	Distress and roughness.
Preventive treatment	4	Periodic, specified by IC, type	1. Scheduled fixed interval 2. Condition low cracking, ravelling	a. Fog seal b. Rejuvena- c. Slurry seal	Life, and (distress not applicable at high distress levels)
Resealing	3	Periodic, specified by IC, type thickness	1. Scheduled fixed interval 2. Condition distress 3. Condition roughness	a. Surface treatment (i.e., chip seal) b. Slurry seal c. Chip seal with shape correction	Resealing Surface type, all distress, roughness (minor).
Overlaying	2	Periodic, specified by IC, type, thickness	1. Scheduled fixed interval 2. Condition roughness	a. Asphalt concrete b. Open-graded cold-mix asphalt c. Auto-level control asphalt concrete	Surface type, distress, rut depth, roughness, strength
Reconstruction	1	Periodic, specified by IC, new pavement	1. Scheduled fixed age 2. Condition roughness	Any surface and base, strength	All pavement characteristics

Source: This study.

of maintenance adequate to assure a normal life for the pavement structure. These components are vital to normal pavement performance, (and, in the case of shoulders, safety aspects) yet they constitute a minor proportion of total maintenance costs. When adequate levels of routine maintenance do not apply, the user should reflect this in the values specified for the pavement strength parameters or deterioration factors. Other items of routine maintenance, e.g., safety installations, signs, vegetation control (other than drainage areas) do not directly affect pavement performance and are left to be determined exogenously. Thus, the user simply specifies a fixed sum per km per year as the basis for costing routine maintenance.

4.3.3 Patching

This operation, usually included under routine maintenance by road authorities because it is an annual operation or recurrent cost, includes mainly surface patching and repair of surfacing distress. Included are skin patches of binder and stone or slurry seal on cracked or ravelled areas, the replacement of the surfacing in small severely-cracked areas, and the filling of potholes. Crack sealing might also be included in this category although it has not been modelled.

The relationships predicting the effect of patching in the submodel represent a mixture of the above forms, as it has not been feasible to distinguish their effects. In the Brazil-UNDP study base, thin slurry seal skin patches were a dominant method of patching, although they were rarely effective on areas of cracking.

The user specifies patching in one of three ways:

1. In a <u>scheduled</u> policy, patching is specified as a fixed maximum area per year (m^2/km), which might, for example, reflect the maximum resources available from the road authority when averaged over all roads within the link category. The amount performed is computed as the lesser of the specified amount and the unpatched severely damaged area, ADAMS. ADAMS is defined as the sum of the areas cracked with wide cracks, ravelled, and potholed at the end of the analysis year before maintenance.

$$\Delta ASP = \min \left[\frac{ASPS_o}{10\ W};\ ADAMS\right]$$

where ΔASP = the area of patching performed, in percent of total carriageway surface area;

$ASPS_o$ = the specified maximum annual patching, in m^2/km; and

$ADAMS = ACRW_b + ARAV_b + APOT_b$; or

2. In one condition-responsive policy, the user may specify the percentage of the severely damaged area, ADAMS, which is to

be patched in each year and impose a limit of the maximum area (m^2/km) per year; or

3. In another condition-responsive policy, the user may specify the percentage of the potholing area to be patched (the area of pothole patching can be computed manually by applying an average depth of patching of 80 mm to a desired patching volume), and a limit of the maximum annual area of patching (m^2/km) per year.

In both cases (2) and (3), the area of patching is given by:

$$\Delta ASP = \min \left[ASPS; \frac{\Delta ASPMAX}{10\ W}\right]$$

where

$$ASPS = \begin{cases} \dfrac{FDAM_o}{100} ADAMS & \text{if responsive to severe damage; or} \\[2ex] \dfrac{FPOT_o}{100} APOT_b & \text{if responsive to potholing only.} \end{cases}$$

$FDAM_o$ = the percentage of the damaged area to be patched in a year, specified by the user;

$FPOT_o$ = the percentage of the potholing area to be patched in a year, specified by the user; and

$\Delta ASPMAX$ = the maximum applicable area of patching in a year, in m^2/km.

The cost of patching per km is computed by multiplying the user-specified unit cost (per m^2) with the area of patching performed, 10 W ΔASP.

When patching is performed, the unpatched damaged area is reduced by the amount of patching. It is assumed that potholing, wide cracking, and ravelling have priorities in that order, and no patching is performed to fix these individual distressed areas until those of higher priorities are completely repaired.

As a result of these incremental changes in pavement condition due to maintenance, the roughness also changes, as predicted by the following relationship, which forms part of the incremental roughness prediction relationship given in Section 4.2.7:

$$\Delta QI_m = \min [0.130 \min (\Delta ASP; 10) + 0.086\ \Delta CRX_m + 4.91\ \Delta APOT_m; 150 - QI_b]$$

where ΔQI_m = the predicted change in road roughness during the analysis year due to maintenance, in QI; and

ΔCRX_m = the predicted change in cracking index (weighted for cracking severity) due to maintenance given by:

$$\Delta CRX_m = 0.62\ \Delta ACRA_m + 0.39\ \Delta ACRW_m.$$

Note that the patching decreases roughness through a reduction of cracked area but usually causes a net slight increase in the roughness under most conditions when no potholes are present. The coefficient represents an average depression or protrusion of about 2 mm for skin patches; when the standard of workmanship observed in maintenance patching differs significantly from this, a specific change could be made in the code (an individual deterioration factor is not available), adjusting the coefficient 0.130 by the ratio of the observed average protrusion/depression to 2 mm. Note also that the effect on roughness is limited to areas of patching not exceeding 10 percent in any year; larger areas of patching are likely to have a net corrective rather than adverse impact on roughness. The patching of potholes is taken to be 90 percent efficient in correcting the roughness due to potholes (i.e., 0.90 0.42 = 0.378 [for m/km IRI] or 4.91 [for QI]).

4.3.4 Preventive Treatments

Preventive treatments are not widespread practice but have been applied in some countries over long periods and are receiving renewed attention because of their low cost and the availability of new asphalt rejuvenating products. Their purpose is to extend the life of bituminous surfaces by retarding the effects of weathering and aging before significant amounts of distress have occurred. The following treatments are included in the submodel.

1. Fog seal is a light sprayed application of bitumen applied on top of a bituminous surface to reduce ravelling by binding the surface stones and to cover oxidized binder with fresh binder. Applications are usually in the range of 0.1 to 0.5 liter/m^2.

2. Rejuvenation is a light application of solvents, oils or plasticizers (e.g., cut-back or fluxed bitumen, emulsified maltenes, etc.) sprayed on to the top of an existing bituminous surface. Dependent upon the effective depth of penetration of the rejuvenator, an oxidized binder is softened towards its original viscosity and becomes less susceptible to cracking and ravelling. Application rates are usually in the range of 0.3 to 0.9 liter/m^2 depending upon dilution.

3. Slurry seal, which is a cold mixture of bitumen emulsion and fine-graded aggregate of 3 to 10 mm maximum size applied in a single layer of approximately one stone-size thickness, can be used as preventive treatment when it is applied as a void-filling coat on surface treatment. By filling the interstices of the surface, the slurry improves durability and retards ravelling, especially when the original binder is

brittle or of inadequate film thickness. As a corrective treatment on asphalt or cracking, slurry tends to be less effective (see Section 4.3.4).

As statistically estimated relationships were not available, tentative relationships have been incorporated in the submodel based on an engineering evaluation of the experience with similar treatments in several countries (Paterson, 1987). In general, preventive treatments are only expected to be effective and economic on relatively low-volume roads where aging effects dominate trafficking effects (for example, less than 1,000 veh/day). The effects of preventive treatments are simulated by:

1. The cracking retardation time, CRT, which is additive to the predicted time of cracking initiation (TYCRA, TYCRW), as illustrated in Figure 4.12. The change in CRT due to preventive treatment maintenance is denoted CRM which has the assumed values for each treatment given in Table 4.13. A maximum limit on the value of CRT due to multiple applications is imposed, defined by CRTMAX and the values given in Table 4.13. The value of CRT, after maintenance, is defined by:

   $$CRT_{(after)} = \min\left[CRT_{(before)} + CRM/YXK;\ CRTMAX/YXK;\ 8\right]$$

   where $YXK = \max(0.1;\ YAX)$.

2. The ravelling retardation factor, RRF, which is multiplicative of the time of ravelling initiation (TYRAV) and a divisor of the rate of ravelling progression ($\Delta ARAV_d$). The change in RRF due to preventive treatment maintenance, RRM, takes the assumed values for respective treatment types given in Table 4.13, subject to a maximum limit, RRFMAX, imposed on RRF for multiple applications. The value of RRF after maintenance is defined by:

   $$RRF_{(after)} = \begin{cases} 1 \text{ if the surface type is AC or OVAC} \\ \min\left[RRF_{(before)}\ RRM;\ RRFMAX\right] \text{ otherwise.} \end{cases}$$

Only one of the treatments can be applied in the analysis year. The treatment policy can be defined either by:

1. A <u>scheduled</u> policy, in which the user specifies a fixed interval between successive treatments, e.g., 3 years, and treatment is applied whenever the surfacing preventive treatment age, AGE1, exceeds this interval; or by

2. A <u>condition-responsive</u> policy, in which treatment is applied at the first signs of cracking or ravelling distress defined by:

   $0 < ACRA_b \leq 15$ or
   $0 < ARAV_b \leq 30$

Table 4.13: Parameters of preventive treatment

Parameters	Pavement type	Parameter values for preventive treatment options		
		Slurry seal	Rejuvenation	Fog seal
CRM	ST on granular base	1.0	3.0	1.6
	OCMS on granular base	0.25	0.75	0.4
	Other combinations	0.5	1.5	0.8
CRTMAX	ST on granular base	6.0	6.0	3.2
	OCMS on granular base	1.5	1.5	0.8
	Other combinations	3.0	3.0	1.6
RRM	All combinations	1.5	1.15	1.3
RRFMAX	All combinations	4.0	2.0	3.0
ΔSN_m	All combinations	0.05	0	0

Note: Values based on available engineering experience, not statistically estimated.
Source: Paterson (1987).

and is constrained by the user-specified limits of the minimum and maximum allowable preventive treatment intervals, in years. Note that treatment is not applied if the distress exceeds either limit above, even if the maximum allowable interval has been exceeded (as may occur in the first analysis year of an old pavement).

If performed, the amount of preventive treatment, in m^2/km, is equal to 1,000 W where W is the width of the carriageway. The cost of preventive treatment per km is computed by multiplying this amount with the user-specified unit cost (per m^2). When preventive treatment maintenance is performed, any surfacing distress (which is minimal) is nullified, and pavement strength is updated by ΔSN_m from Table 4.13, as follows:

$$\Delta[ACRA, ACRW, ARAV, APOT]_m = -[ACRA, ACRW, ARAV, APOT]_b$$

$$AGE1 = 0$$

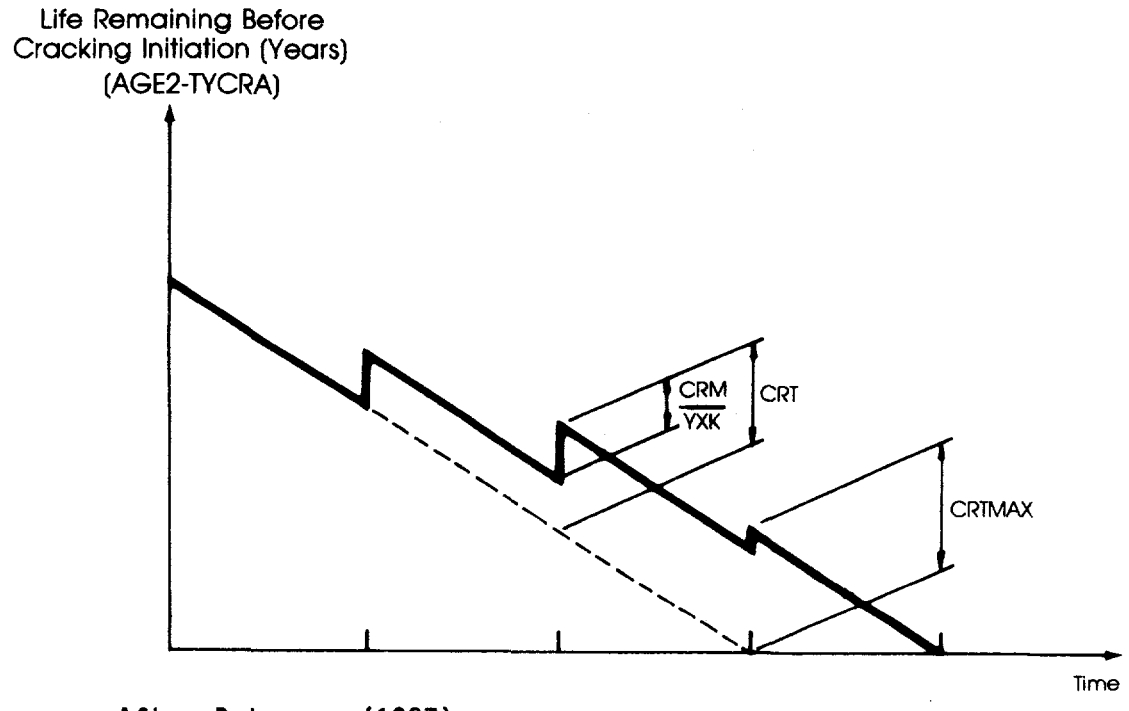
$$SNC_{(after)} = SNC + \Delta SN_m$$

$$DEF_{(after)} = DEF_{(before)} \cdot \left[\frac{SNC_{(after)}}{SNC_{(before)}}\right]^{-1.6}$$

4.3.5 Resealing

Resealing maintenance comprises two thin resurfacing operations which repair surface distress but cause little change to the roughness or structural strength of the pavement; these options are surface treatment (i.e., chip seal) and slurry seal. A third option is surface treatment with shape correction, an alternative in which some reduction of roughness is achieved through the filling of depressions and repair of damaged areas. The corrective material is assumed to be bituminous, with an average thickness of less than 50 mm and placed to a quality of less than that of automatic-levelling paver-finishers. The specification and effects are identical for all options except in the change of roughness and surface type after the application. One unit cost is used for all three but differences can be incorporated through a cost factor. The effects of resealing on subsequent pavement behavior are defined through the reclassification of surface type and adjustments to condition and strength

Figure 4.12: Illustration of effects of preventive treatment on cracking initiation prediction



Source: After Paterson (1987).

variables. In general, slurry seal is not an effective form of reseal on cracked pavements because reflection cracking develops within 2 to 6 months, and the slurry has an unquantified but probably negligible effect on roughness except for possible benefits at high levels (above 60 QI). The user specifies a reseal of fixed type, thickness and strength coefficient to be applied under one of three policies:

1. A scheduled policy, in which a fixed time interval between successive reseals is specified, and the reseal is applied whenever the surfacing age, AGE2, exceeds this interval; or

2. A condition-responsive policy, in which the reseal is applied when the unpatched damaged area, ADAMR, prior to maintenance exceeds a critical level specified by the user, where

   $$ADAMR = ACRA_b + ARAV_b + APOT_b; \text{ or}$$

3. A condition-responsive policy, in which the reseal is applied when the roughness exceeds a critical level specified by the user.

Under any policy, a reseal will not be performed if the surfacing age, AGE2, is less than the user-specified minimum applicable interval, or if the user-specified last applicable year has been exceeded. However a reseal will always be performed if AGE2 exceeds the user-specified maximum allowable interval between reseals.

If performed, the amount of resealing in m^2/km is equal to 1000 W. Under options 1 or 2, if the area of wide cracks is larger than 20 percent, or the area of potholes is not zero at the end of the year before maintenance, preparatory patching of the following amount is assumed to be carried out along with resealing:

$$\Delta ASP = \max\left[0.1\ (ACRW_b - 20);\ 0\right] + APOT_b.$$

Under option 3 preparatory patching is undertaken only of potholes, i.e.,

$$\Delta ASP = APOT_b.$$

The cost of resealing per km is the user-specified unit cost of resealing (per m^2) multiplied by 1000 W. The additional area of preparatory patching is computed as 10 W ΔASP, and the cost is computed by multiplying this area by the unit cost of patching (per m^2). The areas and costs are reported separately under resealing and patching respectively.

Upon resealing the surface type is changed to one of the four types (i.e., RSST, RSAC, SSST or OVSA) as indicated by Table 4.4 depending on the type of reseal (i.e., surface treatment or slurry seal) and the previous surface type. This alters the applicable set of road deterioration relationships, as described above, to represent different deterioration behavior after performing the reseal.

The maintenance effect of resealing is to set the areas of cracking (for the new surfacing layers only), ravelling and potholing to zero, as follows:

$$\Delta[ACRA, ACRW, ARAV, APOT]_m = -[ACRA, ACRW, ARAV, APOT]_b$$

The effects on roughness (QIS) of the first two options are not well-determined (Paterson, 1987). Data from a number of sources indicate step changes in roughness due to resealing which range from small negative to small positive increments. Currently, it is thought that the positive increments represent either the transient effects of the early life of a surface treatment before stone embedment occurs, or the effects of roughness measurement error (which can be large on surface treatments). Thus, pending future research, positive increments were excluded, and the following relationship was adopted in the submodel, as illustrated in Figure 4.13(a). It is given that patching has repaired all potholes with an effect QIP.

$$\Delta QI_m = \min [0.086\ \Delta CRX_m + QIP + QIS;\ 150 - QI_b]$$

where $$QIP = \max (4.91\ W\ \Delta APOT_m;\ -60)$$

$$QIS = \begin{cases} \min\{0;\ \max [0.3\ (70 - QI_b - QIP);\ -6]\} \\ \quad \text{if reseal with surface treatment} \\ \min\{0;\ \max [0.3\ (60 - QI_b - QIP);\ -1.2\ H_o]\} \\ \quad \text{if reseal with slurry seal} \end{cases}$$

For reseal with shape correction, the roughness change is given by the following, as illustrated in Fig. 4.13(b):

$$\Delta QI_m = \min \{0,\ \max [-0.0075\ H_{sc} QI'_b,\ -0.0225\ H_{sc} \max (QI'_b - 52,\ 0)]\}$$

where H_{sc} = average thickness of reseal including the shape correction layer, in mm,

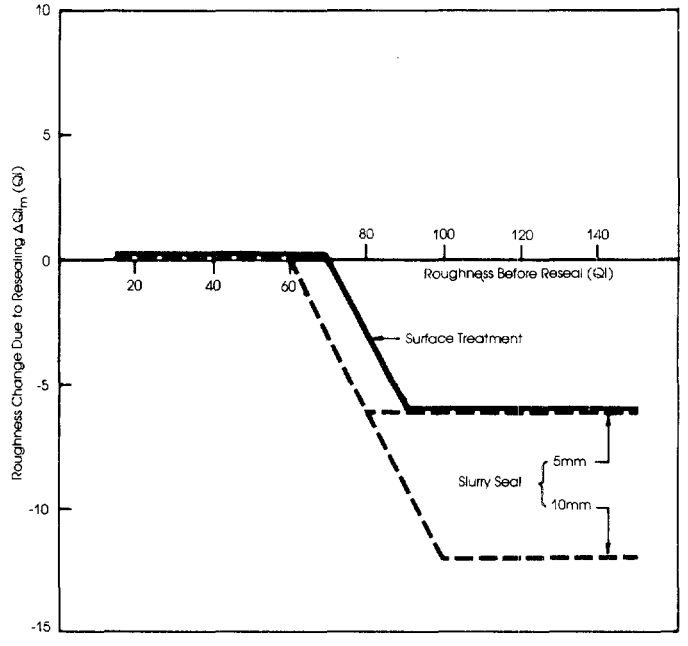
$= \min (H_o,\ 50)$; and

$QI'_b = QI_b + QIP$.

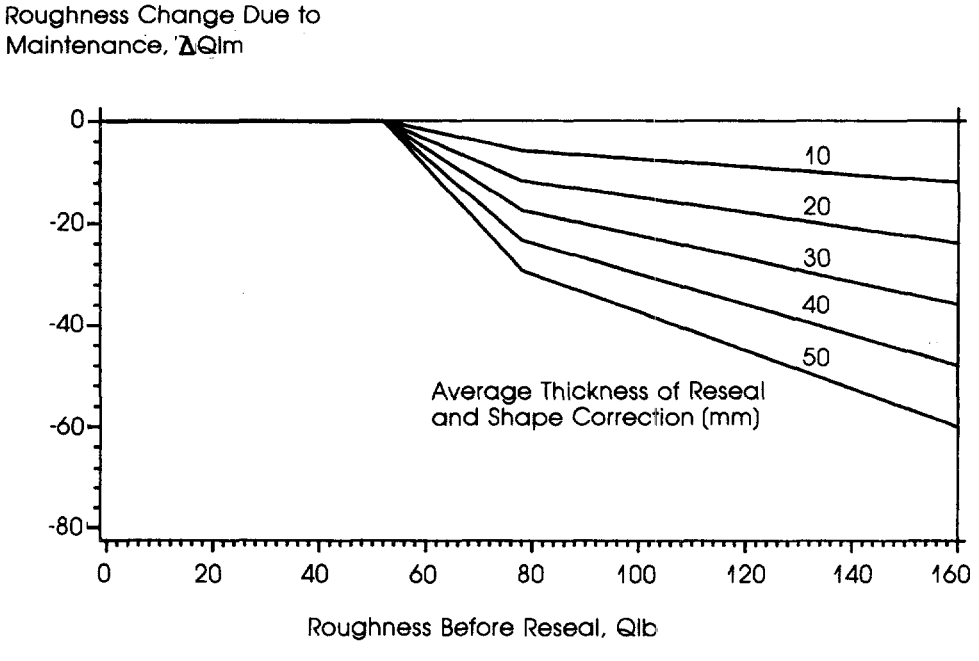
The cracking retardation time, ravelling retardation factor and preventive treatment and surfacing ages are reinitialized to mark the beginning of the new deterioration phase. The resealing operation is assumed to be performed under good quality control, so the construction fault code, CQ, is set to zero endogenously. To take account of the net strengthening of the pavement due to both maintenance and cracking, the pavement strength parameters are updated as follows:

Figure 4.13: Maintenance effect of resealing on roughness

(a) Surface treatment and Slurry seal options



(b) Surface treatment with shape correction option



Source: This study.

$$SNC_{(after)} = \max\,[1.5;\ SNC_{(before)} + 0.0394\ a_o\ H_o - \Delta SNK]$$

$$DEF_{(after)} = DEF_{(before)}\ \left[\frac{SNC_{(after)}}{SNC_{(before)}}\right]^{-1.6}$$

where a_o = the strength coefficient of the reseal;
H_o = the thickness of the reseal, in mm; and
ΔSNK = the predicted reduction in the structural number due to cracking since the last pavement reseal, overlay or reconstruction (see Section 4.2.7).

In updating the thickness parameters the reseal thickness becomes the new surfacing thickness, HSNEW, and the total thickness of previous surfacings becomes HSOLD. The areas of previous cracking (PCRA, PCRW) are updated to equal the cracking in the current surfacing before resealing, and a weighting, w, of the cracking in the previous surfacing, as given by:

$$PCRA_{(after)} = \begin{cases} ACRA_b & \text{if } ACRA_b \geq PCRA_{(before)} \\ w\ ACRA_b + (1 - w)\ PCRA_{(before)} & \text{if } ACRA_b < PCRA_{(before)} \end{cases}$$

and

$$PCRW_{(after)} = \begin{cases} ACRW_b & \text{if } ACRW_b \geq PCRW_{(before)} \\ w\ ACRW_b + (1 - w)\ PCRW_{(before)} & \text{if } ACRW_b < PCRW_{(before)} \end{cases}$$

where w = the weight used for averaging the cracking in the old and new surfacing layers, as given by the assumed relationship:

$$w = \min\,(0.70 + 0.1\ H_o;\ 1).$$

4.3.6 Overlay

The overlay operation in the submodel applies only to bituminous overlays placed by mechanical paver-finisher; other overlays are dealt with as described below. Three options are provided:

1. Open-graded cold-mix asphalt, regular or manual levelling control;

2. Hot-mixed asphalt concrete, regular or manual levelling control; and

3. Hot-mix asphalt concrete, with long-base automatic levelling (control (base longer than 5m).

The overlay operation in the model applies to single-layer overlays, but double-layer asphaltic overlays of less than 125 mm thickness may be specified under this operation by regarding the two layers as one compound layer. Double-layer overlays, such as granular overlay with bituminous surfacing, or cemented layers, or asphaltic overlays of 125 mm thickness or greater, are best specified under the reconstruction operation. In that case, the overlay is modelled as a new pavement with the top layer as the new surfacing and the lower layer as the new base.

The effects of overlay on pavement deterioration are defined through the reclassification of surface type and the associated deterioration prediction relationships.

An overlay policy comprises an overlay of fixed type, thickness and strength coefficient specified by the user and applied as follows:

1. Scheduled policy, in which the overlay of fixed specifications is applied whenever the construction (or 'rehabilitation') age, AGE3, equals or exceeds a fixed time interval specified by the user, or

2. Condition-responsive, in which the overlay of fixed specifications is applied when the roughness before maintenance exceeds a maximum allowable roughness specified by the user.

Under either policy, an overlay is not performed if the construction age, AGE3, is less than the user-specified minimum applicable overlay interval or if the last applicable year has been exceeded. An overlay is also not performed if either the preventive treatment age, AGE1, or the surfacing age, AGE2, is less than the respective minimum preventive treatment or reseal intervals. An overlay is performed, however, when the construction age, AGE3, exceeds the (optional) maximum allowable overlay interval specified by the user.

If performed, the amount of overlay, in m^2/km, is equal to 1,000 W. The cost of the overlay per km is computed by multiplying with this amount the user-specified unit cost (per m^2).

When an overlay is performed, the surface type is changed to either of the two types: asphalt concrete (OVSA) or cold mix (OCMS) as indicated by Table 4.4, depending on the type of overlay specified but regardless of the old surface type. This alters the applicable set of surfacing distress relationships, as described above, to represent different deterioration behavior after performing this maintenance operation.

For the immediate effect of overlay, the model sets to zero the amounts of cracking (for the new surfacing layers only), ravelling, and potholing and the cracking retardation time, the construction fault code; and resets the ravelling retardation factor, the preventive treatment, surfacing and construction ages. The rut depth is reduced by 85 percent, in recognition of the residual remaining after the compaction of a single

layer overlay. The roughness is reset to the (optional) value specified by the user, QII_o, or to the default values dependent upon the quality of paving, as computed in the formulas for ΔQI_m below, and illustrated in Figure 4.14:

Option (a): Regular paver

$$\Delta QI_m = \begin{cases} QII_o - QI_b & \text{if } QII_o \text{ is provided} \\ \min\{0;\ 20 + 10\,k_{ov} + [28 \max(QI_b' - 50;\ 0)/\max(H_o';\ 28)] - QI_b'\} & \text{otherwise.} \end{cases}$$

Option (b): Automatic-levelling long-base paver

$$\Delta QI_m = \begin{cases} QII_o - QI_b & \text{if } QII_o \text{ is provided} \\ \min\{0;\ \max[(19.42 - 0.78\,QI_b' - 0.068\,H_o); \\ \qquad (-19.5 - 0.008\,QI_b' \max(H_o - 20;\ 0))]\} & \text{otherwise;} \end{cases}$$

where

QII_o = the initial road roughness after overlay specified by the user as an option, in QI:

QI_b' = max (13; QI_b + 4.91 APOTm);

H_o = thickness of the overlay, in mm.

Ho' = min (Ho; 80); and

H_{ov} = 3 - [min (HO; 80) + min (Ho; 40)]/40.

In addition the pavement strength parameters are updated to take account of the net change in pavement strength due to the new overlay and the underlying cracks (if any). The procedure for the modification is the same as that for reseal described in Section 4.3.4, except for the following redefinitions:

a_o = the strength coefficient of the overlay; and
H_o = as defined above; and

$$w = \begin{cases} \max(HSNEW_{(before)}/HSOLD_{(after)};\ 0.6) & \text{if base type is not cemented} \\ \max[HSNEW_{(before)}/(HSOLD_{(after)} + HBASE);\ 0.6] & \text{otherwise.} \end{cases}$$

Double-layer overlay with bituminous second layer can be treated in the same way as single layer overlays with adaptation of the input values as follows. The thickness and the strength coefficient of overlay in this case, are the sum and the weighted average of those two layers, respectively. The type of the overlay (asphalt concrete or cold mix) is determined by the top layer of the overlay.

4.3.7 Pavement Reconstruction

Pavement reconstruction applies in the model to all works that require re-specification of the surfacing and base types, and pavement thicknesses and strength parameters. This encompasses strengthening by multiple-layer overlays thicker than 125 mm, recycling of the base and/or surfacing layers, and membrane-interlayer overlays. But it excludes widening, road realignment, and other geometric reshaping, which should be specified as new construction through the road construction submodel (Series B). All surface types and base types and combinations thereof listed in Table 4.4 are permitted except for surface type RSAC (reseal on asphalt concrete).

The reconstruction operation is defined by user-provided specifications for new surfacing and base layer thicknesses and material types, the construction quality code for surface treatments, resilient modulus for cemented base layers, and either the increment in structural number or the final structural number after reconstruction. (In order to specify a membrane interlayer overlay, in cases when the membrane is known to be effective in preventing crack reflection, the user can simply specify the new surfacing type, thickness and strength coefficient and redefine the base type if considered appropriate. The membrane need not be specified explicitly, as the reconstruction operation automatically resets all history and condition variables to zero.) The reconstruction policy may be either:

1. Scheduled, in which the fixed reconstruction is to be performed whenever the construction age, AGE3, equals or exceeds the user-specified maximum allowable age or

2. Condition-responsive, in which the fixed reconstruction is performed when the roughness before maintenance, QI_b, equals or exceeds the user-specified maximum allowable roughness.

Again under either policy, reconstruction is not performed if the construction age, AGE3, is less than the minimum reconstruction interval, nor if the last applicable year has been exceeded, nor if the preventive treatment and surfacing ages are less than their respective minimum intervals, nor if it is already a construction opening year. Reconstruction is always performed if the construction age, AGE3, exceeds the maximum allowable reconstruction interval, if specified by the user.

If performed, the amount of pavement reconstruction, in m^2/km, is equal to 1,000 W. The cost per km is computed by multiplying the user-specified unit cost (per m^2) with this amount.

When pavement reconstruction is performed, the surface and base types are converted to the new types specified by the user, which can be identical to the old types. The model sets to zero the amounts of cracking (for both new and old surfacing layers), ravelling, potholing and rut depth; resets the preventive treatment, surfacing and construction ages and the cracking retardation time and ravelling retardation factor; and sets the roughness to the (optional) value specified by the user or the default value for the given surface type (see QII_0 in Section 4.2.1). In addition, the construction quality code is set to the (optional) value specified by the user or the default value of zero.

The pavement strength parameters are updated as follows:

$$SNC_{(after)} = \begin{cases} SN_0 + SNSG & \text{if the user specified a new structural number} \\ SNC_{(before)} + \Delta SN & \text{if the user specified an increase in the structural number} \end{cases}$$

$$DEF_{(after)} = \begin{cases} 6.5\ SNC_{(after)}^{-1.6} & \text{if new base is not cemented} \\ 3.5\ SNC_{(after)}^{-1.6} & \text{if new base is cemented} \end{cases}$$

where SN_0 = the user-specified structural number for the reconstructed pavement; and

ΔSN = the user-specified increment in the structural number due to the pavement reconstruction.

4.3.8 Pavement Parameters after Maintenance

Following the logic set out in Section 4.1.2, the submodel computes the pavement condition for the beginning of the next analysis year by adding the maintenance effects to the condition before maintenance as follows:

$$[CONDITION]_{a(next\ year)} = [CONDITION]_b + \Delta[CONDITION]_m$$

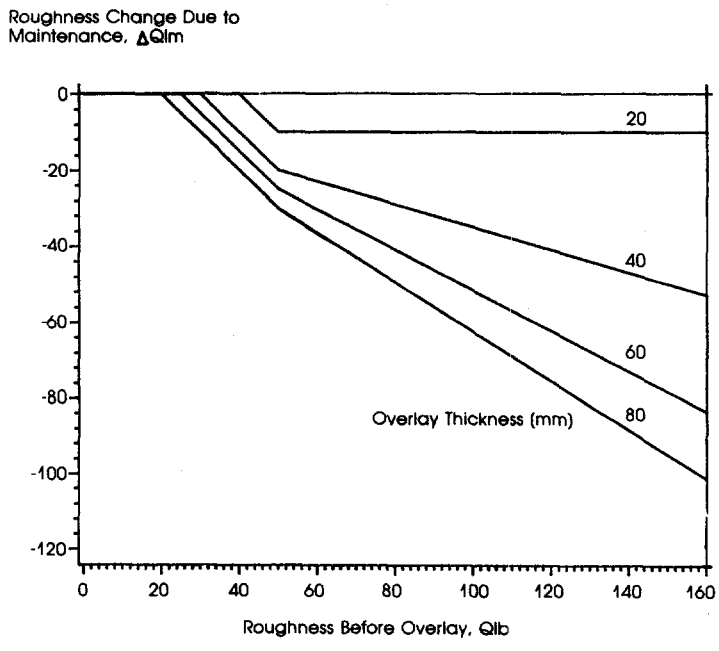
A constraint is placed on the post-maintenance roughness so that it is not lower than a practical minimum of 15 QI or 1.2 m/km IRI, i.e.,

$$QI_{a(next\ year)} = \max\ (QI_b + \Delta QI_m;\ 15).$$

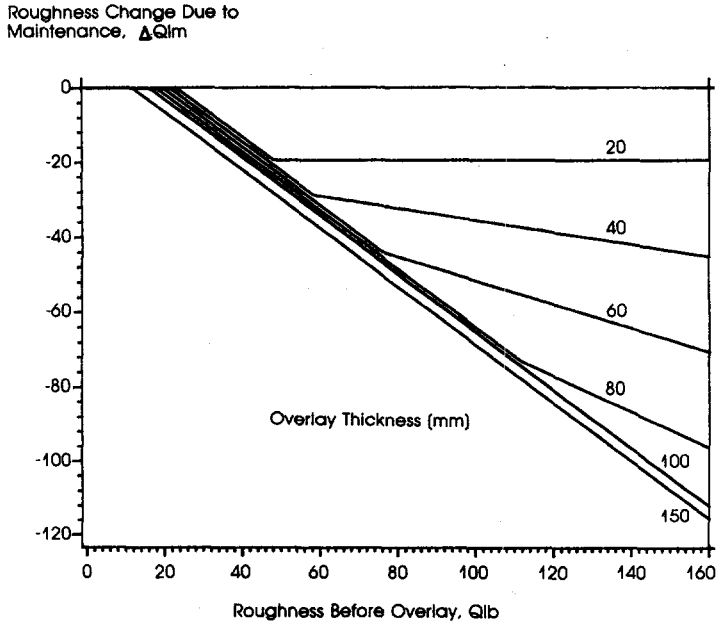
As HDM-III uses the section as the smallest unit of road for computing vehicle operating costs, the average road roughness of the section over the analysis year is required, which is computed as the arithmetic average of the corresponding averages for the three subsections of weak, medium and strong behavior. The average is given by:

Figure 4.14: Maintenance effect of asphalt overlays on roughness

(a) Regular and manual control paver-finishers



(b) Long-base automatic-levelling paver-finishers



Source: This study.

$$QI_{avg} = \sum_{j=1}^{3} (QI_{aj} + QI_{bj})/6$$

where QI_{aj} and QI_{bj} denote the roughness of each subsection j at the beginning and the end of the analysis year before the road maintenance decision, respectively.

4.4 UNPAVED ROAD LOGIC

4.4.1 Classification, Concepts, and Logic

Unpaved roads comprise the lower classes of the road network hierarchy, and generally carry low volumes of traffic ranging from a few vehicles to up to several hundred vehicles per day. The geometric standards vary considerably and it is necessary to make a primary classification of unpaved roads into engineered roads, which have controlled alignment, formation width, cross-section profile and drainage; and tracks, which are essentially ways formed by trafficking along natural contours with or without the removal of topsoil. Unpaved roads classified in a country's network are usually engineered or partly engineered roads, and tracks are usually not classified.

The analysis of deterioration and maintenance effects in this submodel is designed primarily for engineered unpaved roads, of either gravel or earth surfacing, because the empirical models are based on a variety of such roads. When necessary it is possible to use the relationship also for tracks as a first estimate, but the user needs to be aware that the environmental effects of drainage and rainfall may be poorly represented.

The deterioration of unpaved roads is characterized primarily by roughness and by material loss from the surfacing. The prediction relationships for these are based on analyses of the Brazil-UNDP study (Visser, 1981; Paterson, 1987). Wheel path ruts also develop under traffic but the ruts are usually devious or poorly defined and often mixed with water-induced surface erosion. Thus the concept of rut depth is not used in HDM-III and is subsumed in the property of roughness; prediction relationships may be found in Visser (1981). The looseness of surfacing material, which was analyzed in the Kenya study (Hodges et al., 1975), was also observed in the Brazil-UNDP study (GEIPOT, 1982) but as it was found to have no substantial effect on vehicle speed, no prediction relationships have been incorporated in HDM-III. Finally, road passability is an important criterion for upgrading tracks or earth roads to gravel roads - provision is made in the model for an increase in vehicle operating costs (by a factor specified by the user, reflecting the economic effects of reduced passability when the gravel thickness drops below a minimum level - this is discussed in Sections 4.5.3 and 5.3.7.

The maintenance of unpaved roads comprises:

1. Periodic grading by motorized or towed grader to restore surfacing gravel from the shoulders to the roadway and to reduce roughness;

2. Spot regravelling to repair potholes;

3. Gravel resurfacing to replace or augment the gravel surfacing layer in response to material loss; and

4. Routine-miscellaneous maintenance of drainage channels and verges.

The periodic grading of unpaved roads is usually undertaken on a more-or-less regular basis for management purposes, either seasonally or frequently enough to keep the roughness within tolerable limits. The 'routine' tasks 2 and 4 above are often achieved at the same time as 1.

These repeated cycles of roughness deterioration and grading maintenance are treated as continual by the submodel. The average roughness during each analysis year is computed as a function of the roughness at the beginning of the year, of material, traffic, geometry and rainfall parameters and the specified grading frequency. Over a period of time depending on the traffic volume and frequency of grading, the annual average roughness tends towards a long-term average roughness which is also computed.

Maintenance of the gravel surfacing is accounted each analysis year through the surfacing thickness and the net change from material loss, spot regravelling and gravel resurfacing maintenance. The material loss from earth roads, although computed, is accounted only for the purpose of predicting "spot regravelling" quantities and is otherwise ignored.

The computational logic described above is illustrated by the flow diagram in Figure 4.15 and detailed in Section 4.4.6. In order to simplify the logic, an unpaved road is considered to comprise two layers, a gravel surfacing and a subgrade. A gravel road has both layers, but an earth road has a zero thickness of gravel surfacing and its surface characteristics are those of the subgrade. When a gravel road loses all of its gravel surfacing, then its classification reverts to that of earth road. Upon gravel resurfacing, all unpaved roads become gravel roads by definition of the new surfacing layer.

Deterioration is predicted using the properties of the surfacing layer, whether that be "gravel" or "subgrade," as it is defined for the analysis year. Thus the user must specify the physical properties of both gravel surfacing and subgrade for unpaved roads.

4.4.2 Material Properties

Previously, deterioration relationships have been categorized by material type (lateritic, quartzitic, coral, volcanic, etc.), but from the Brazil-UNDP study it has been possible to replace these by material properties which should improve the transferability of the relationships.

The material properties which were found to affect the rate of deterioration in Brazil include the maximum particle size, the particle size distribution and the soil plasticity (Paterson, 1987). The specific soil properties, which are required inputs for HDM-III as defined below, are used subsequently to define various summary numerics of the particle size distribution which are parameters in the deterioration prediction equations. The minimum and maximum levels of roughness, QIMIN and QIMAX, are predicted endogenously from the soil properties but the user may override those by specifying input values. The soil properties are defined for both the 'gravel' and 'subgrade' layers (as described in Section 4.4.1) denoted by the subscript j, where j = g for gravel surfacing layer, and j = s for subgrade (or earth road surfacing) layer, in Table 4.14.

4.4.3 Traffic Loading Measures

The traffic loading variables used in predicting unpaved road deterioration are simply those of two-way traffic counts for all vehicles and for light (ADL) and heavy (ADH) vehicles, as defined in Table 4.14. Note that Table 5.1 provides default classifications for light and heavy vehicles. The variable ADT, which equals ADL + ADH, is used in predicting material loss, and the variables ADL and ADH in predicting roughness.

4.4.4 Road Geometry Measures

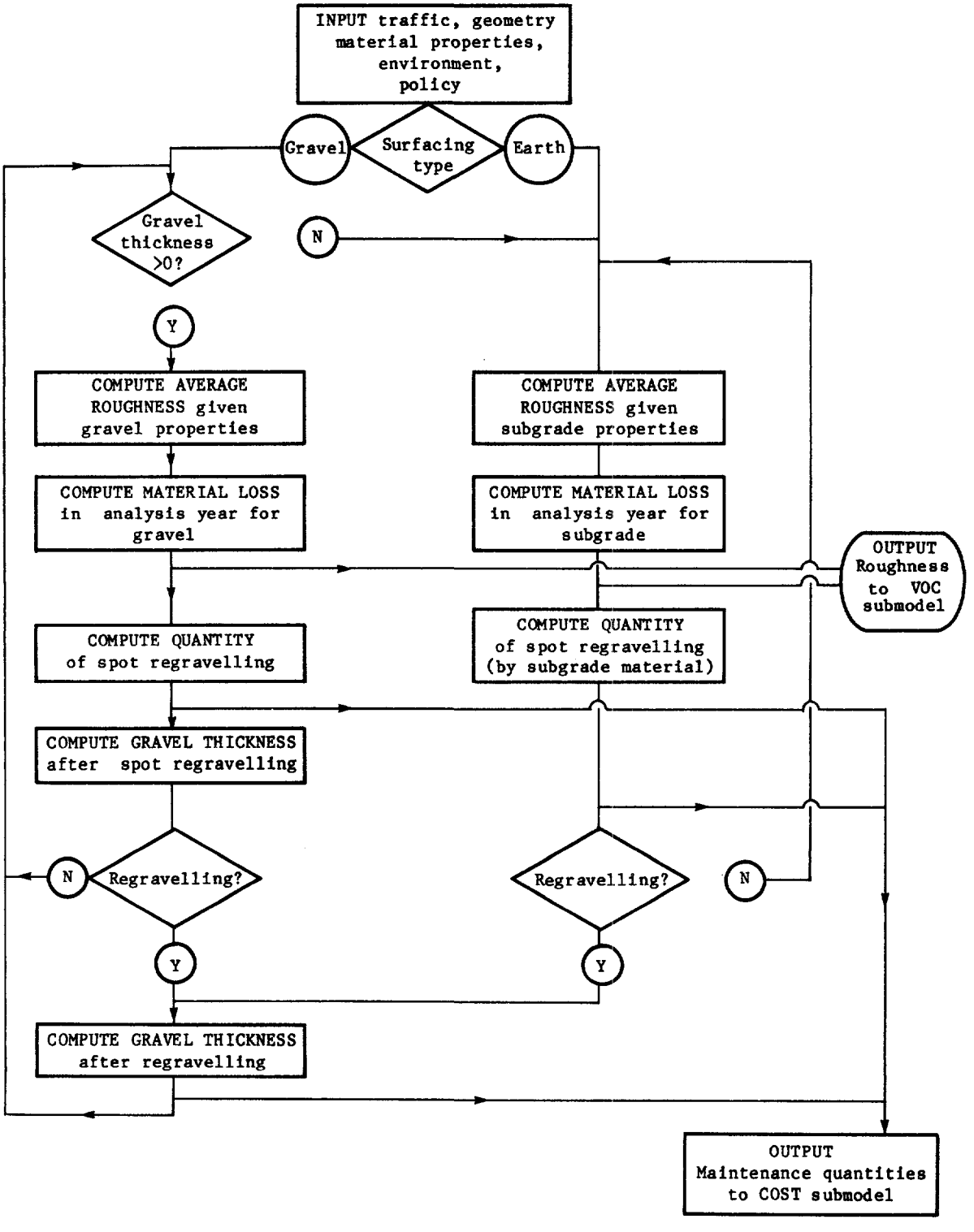
The geometric characteristics found to influence the deterioration of unpaved roads in the Brazil-UNDP study were horizontal curvature (c) and longitudinal gradient (here represented by the rise plus fall variable, RF). Roughness progression, and in particular the maximum roughness, is influenced by both characteristics. In material loss prediction the horizontal curvature affects the rate of traffic-induced material whip-off and the gradient interacts with rainfall in causing erosion. Cross-sectional geometry, including crown, camber and superelevation, were not measured in the study and are discussed in the following section. The average shoulder width (WS) is used to compute the amount of gravel used in spot regravelling and gravel resurfacing. The variables RF, C and WS are defined in Table 4.14.

4.4.5 Environment: Climate and Drainage

While the climate of the Brazil-UNDP study area is classed as humid to warm- or wet-humid, the rainfall pattern was seasonal, ranging from precipitations of less than 20 mm per month and air humidity less than 40 percent during a continuous six to eight months of a year, to precipitations of 200 to 600 mm per month and air humidity in excess of 60 percent over four months of a year. The effects of the full range of highly seasonal rainfall were analyzed in the study, and are represented by the average monthly rainfall in the deterioration prediction relationships. The predictions of annual average roughness and material loss transform this to an annual average rainfall and thus make no specific distinction between uniform- and seasonal-rainfall climates

Geometric cross-sectional characteristics, particularly crown, camber, table side-drains and run-off points, have pronounced effects on

Figure 4.15: Logic sequence of road deterioration and maintenance submodel: unpaved roads



Source: This study.

Table 4.14: Definition of primary variables for unpaved roads

Variable		Definition
ADH	=	the average daily heavy vehicle traffic (GVW >= 3,500 kg) in both directions, in vehicles/day.
ADL	=	the average daily light vehicle traffic (GVW < 3,500 kg) in both directions, in vehicles/day; and
ADT	=	the average daily vehicular traffic in both directions, in vehicles/day;
C	=	the average horizontal curvature of the road, in degrees/km (as defined in Figure 5.1);
$D95_j$	=	the maximum particle size of the material, defined as the equivalent sieve opening through which 95 percent of the material passes, in mm;
MG_j	=	slope of mean material gradation, as defined in Section 4.5.1;
MGD_j	=	dust ratio of material gradation, as defined in Section 4.5.1;
PI_j	=	the plasticity index of the material, in percent;
$P075_j$	=	the amount of material passing the 0.075 mm sieve (or ASTM No. 200 sieve), in percent by mass;
$P425_j$	=	the amount of material passing the 0.425 mm sieve (or ASTM No. 40 sieve), in percent by mass;
$P02_j$	=	the amount of material passing the 2.0 mm sieve (or ASTM No. 10), in percent by mass;
QI_{avg}	=	average roughness during analysis year, in QI;
$QI_{(after)}$	=	roughness after grading, in QI;
$QI_{(before)}$	=	roughness before grading, in QI;
$QIMIN_j$	=	the minimum roughness of the material (either estimated in Section 4.5.1 or specified), in QI;
$QIMAX_j$	=	the maximum roughness of the material (either estimated in Section 4.5.1 or specified), in QI;
RF	=	the average absolute rise plus fall of the road, in m/km (as defined in Figure 5.1) (note: RF = 10 times average absolute gradient in percent); and
WS	=	average width of shoulder, in meters;

drainage and deterioration during high rainfall. In the study area, roughness levels on level, tangent sections that were poorly drained were very high during wet periods due largely to the rapid development of potholes. On vertical grades, roughness levels were frequently low despite extensive erosion by surface run-off because the longitudinal profile was affected less than the transverse profile. The study sections generally had moderate drainage facilities and maintenance, and positive crowns. The prediction relations therefore apply to unpaved roads with moderate to good cross-sectional geometry and for dry to wet conditions but may not apply to 'bathtub' type roads with negative crown or lack of surface drainage in high rainfall conditions.

4.4.6 Basic Computational Procedure

The model assumes that the grading operations and spot regravelling specified for each year, both for gravel and earth roads, are distributed uniformly throughout the year. However, the gravel resurfacing operation, when it occurs, is assumed to be carried out at the end of the year. Like the periodic paved road maintenance operations, gravel resurfacing is not permitted in an effective construction completion year. The computational procedure for road deterioration and maintenance of the unpaved roads for each analysis year follows the flow diagram of Figure 4.15 and comprises the following steps:

1. Initialize road characteristics and traffic loading variables at the beginning of the analysis year.

2. If earth road skip to step 3. Otherwise, check whether the gravel thickness is zero (i.e., no gravel remaining) at the beginning of the analysis year. If the thickness is zero, reset the road type to earth.

3. If grading is specified compute the annual average road roughness as a function of the grading frequency, traffic volume, environmental conditions, and attributes of the gravel (if gravel road) or the subgrade (if earth road). Otherwise, if no grading is specified, set the average roughness equal to the predicted maximum roughness (Section 4.5.1).

4. Compute the depth of material loss during the analysis year as a function of the traffic volume, monthly rainfall, and road geometry and the attributes of the gravel (if gravel road) or the subgrade (if earth road) (Section 4.5.2). Compute the average thickness of spot regravelling to be provided during the analysis year (Section 4.5.3).

5. If earth road skip to step 6. Otherwise, compute the gravel thickness at the end of the analysis year before gravel resurfacing decision. This is equal to the thickness at the beginning of the year minus the depth of gravel loss plus the spot regravelling thickness. If the gravel thickness is negative, reset it to zero.

6. Determine whether gravel resurfacing is to be carried out (at the end of the analysis year) according to the criterion specified by the user as described below. If gravel resurfacing is performed set the road type to gravel road regardless of the previous type and compute the gravel thickness at the end of the analysis year after resurfacing as equal to the thickness before resurfacing plus the resurfacing thickness.

7. Compute the corresponding road maintenance quantities and costs by operation.

8. Store the results for later use in the vehicle operating cost submodel and in the evaluation and reporting phase.

4.4.7 Initialization of Variables

At the beginning of the analysis year the traffic variables are computed based on the user-specified traffic data (Series E). The values of the environment, road geometry, and material property variables are provided in one of three ways, in the same manner as the similar variables for paved roads, that is:

1. From the preceding analysis year, if the analysis year is neither the first year of the analysis period nor a construction opening year;

2. From the existing link characteristics data (Series A) if the analysis year is the first year of the analysis period; or

3. From the construction option data (Series B) if the analysis year is a construction opening year.

The only history variable for unpaved roads is the gravel age, denoted by GAGE, which is relevant only for gravel roads. It is defined as the number of years elapsed since the latest gravel surfacing or resurfacing and is initialized as follows:

1. When the analysis year is not a construction opening year, the value of GAGE is provided either (a) from the preceding year (if the analysis year is the second or a subsequent year of the analysis period) or (b) from the existing link characteristics data (if the analysis is the first year of the analysis period), and is increased by one year; and

2. When the analysis year is a construction opening year (of a gravel road project), the value of GAGE is set to one and the unpaved road surface type to gravel, irrespective of the previous surface type.

4.5 UNPAVED ROAD DETERIORATION AND MAINTENANCE

4.5.1 Road Roughness

The roughness of unpaved roads increases through the shear, mechanical disintegration, and erosion of the surfacing material caused by traffic and surface water runoff. Roughness levels are usually 4 to 15 m/km IRI (50 to 200 QI) although lower levels sometimes occur with fine materials. Roughness in excess of 13 m/km IRI (180 QI) is usually related to depressions, potholes or transverse erosion gullies, and levels above 22 m/km IRI (300 QI), which correspond to numerous wheel-sized potholes, are very rare and usually apply only on short sections or unclassified tracks. The roughness modelled for economic evaluation is the profile in the wheelpaths of the traffic, since this generates the vehicle operating costs. The location of the wheelpaths tends to vary when roughness reaches high levels as vehicles seek to minimize the dynamic impact, hence the prediction of roughness progression must take this self-regulating tendency into account. On account of the high variability of material properties, drainage, surface erosion and the high roughness levels of unpaved roads, prediction errors tend to be large, in the order of 1.5 to 2.5 m/km IRI (20 to 32 QI) standard error, or equivalent to 95 percentile confidence intervals of 20 to 40 percent.

A number of different model forms have been applied to roughness progression and to the effects of maintenance grading (Hodges et al., 1975; Visser, 1981; Paterson, 1987). As the objective of policy analyses can be satisfied by computation of the average roughness resulting from a specifiedpolicy, the model selected for predicting roughness was one which both represented the progression and grading phases of the roughness cycle realistically and also permitted a closed-form solution. The model form and its derivation are described in detail elsewhere (Paterson, 1987). The primary principles and parameter estimates are as follows.

Although the IRI roughness measure, or other compatible measure, could have been used in the following relationships because many parameters are dimensionless, we have kept the nomenclature of QI for roughness for internal consistency with the remainder of the model.

Roughness progression

In previous models, progression followed either cubic (Hodges et al., 1975) or exponential (Visser, 1981; Paterson, 1987) concave curves which, unless restrained, led to unrealistically high predictions of roughness for policies of infrequent grading. The model form adopted here constrains the roughness to a high upper limit, or maximum roughness ($QIMAX_j$), by a convex function in which the rate of progression decreases linearly with roughness to zero at QIMAX conforms well with practical observations. The predictions of both forms differ significantly only at high levels of roughness; at low levels of roughness the concave curve is often more realistic in shape, but quantitatively there is little difference between the two. From the Brazil-UNDP study, the maximum roughness was found to be a function of material properties and road geometry, and the rate of roughness progression to be a function of the

roughness, maximum roughness, time, light and heavy vehicle passes and material properties, as illustrated in Figure 4.16 and given by (Paterson, 1987):

$$QI\ (TG_2) = QIMAX_j - b\ [QIMAX_j - QI\ (TG_1)]$$

where $QI\ (TG_1)$ = roughness at time TG_1, in QI;

$QI\ (TG_2)$ = roughness at time TG_2, in QI;

TG_1, TG_2 = time elapsed since latest grading, in days;

$b = \exp\ [c\ (TG_2 - TG_1)]$; where $0 < b < 1$;

$c = -\ 0.001\ (0.461 + .0174\ ADL + .0114\ ADH - 0.0287\ ADT\ MMP)$;

$QIMAX_j = \max\ [279 - 421\ (0.5 - MGD_j)^2 + 0.220\ C - 9.93\ RF\ MMP;\ 150]$;

MGD_j = material gradation dust ratio, defined as

$$MGD_j = \begin{cases} 1 \text{ if } P425_j = 0 \\ PO75_j/P425_j \text{ if } P425_j > 0; \text{ and other} \end{cases}$$

variables = as defined previously in Section 4.4.3 - 4.4.5.

Note: The standard error of this prediction on the original data base was 19.8 QI (1.5 m/km IRI).

Effect of compaction on roughness progression

Observations on gravel and earth roads in the first few grading cycles after construction or rehabilitation with full mechanical shaping and compaction, indicate rates of roughness progression that are much slower than given by the model above, which was derived from roads under repeated grading cycles with no special compaction (Paterson, 1987). Thus if "mechanical compaction" is specified in the model inputs, the coefficient c is reduced, initially to one quarter of its predicted value and rising to the full predicted value after a few grading cycles, but in a period not exceeding 4 years, as follows

$$c' = c \min\ [1,\ 0.25\ t \max\ (1,\ n^{0.33})]$$

where

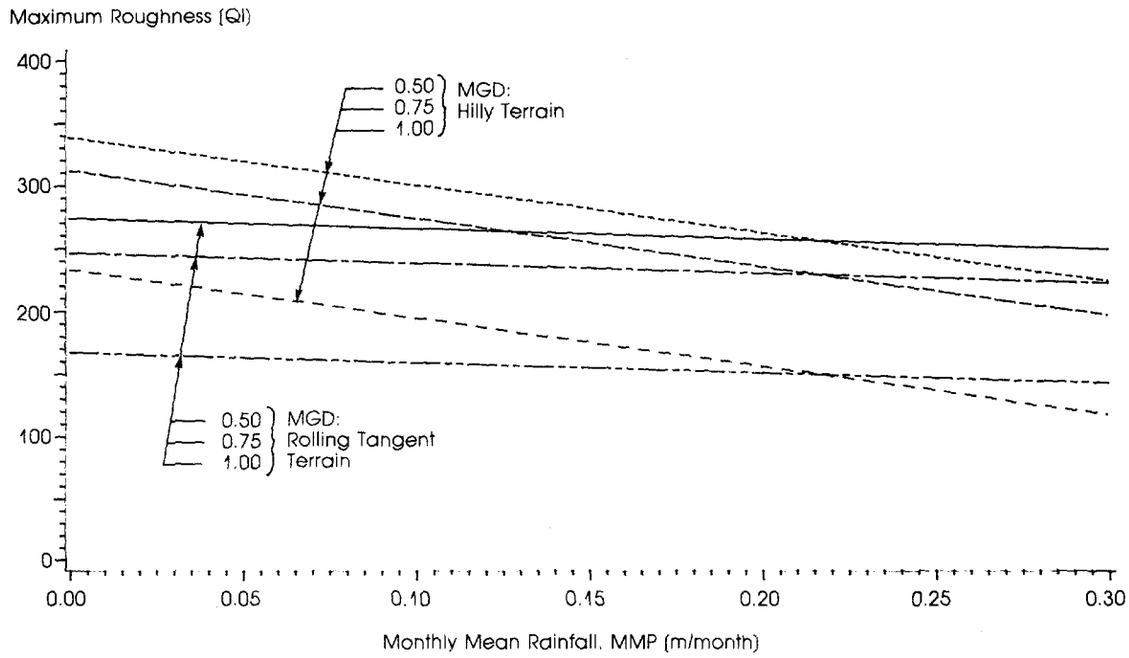
t = time in years since regravelling or construction with mechanical compaction

n = frequency of grading, cycles/year.

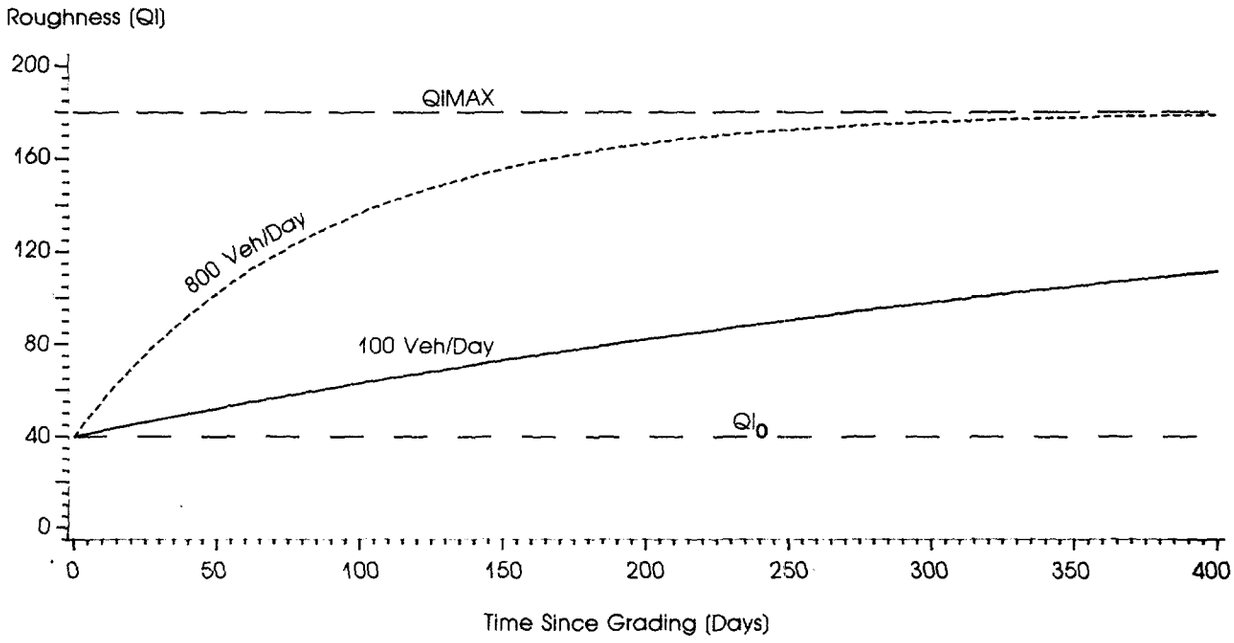
and thus

Figure 4.16: Predictions of maximum roughness and roughness progression for unpaved roads

(a) Estimated maximum roughness for the two geometrics and three materials



(b) Roughness progression for two traffic volumes



Source: Paterson (1987).

$$b' = \exp(c' \, 365/n)$$

where b', c' are the values of b and c above when mechanical compaction is in effect.

Effect of grading

The effect of grading maintenance on roughness was found to depend on the roughness before grading, the material properties and the minimum roughness ($QIMIN_j$) (Paterson, 1987). The minimum roughness, below which grading cannot reduce roughness, increases as the maximum particle size increases and the gradation of the surfacing material worsens. The prediction of roughness after grading is expressed as a linear function of the roughness before grading, dust ratio and the minimum roughness, as illustrated in Figure 4.17 and given by:

$$QI_{(after)} = QIMIN_j + a\,[QI_{(before)} - QIMIN_j]$$

where $QI_{(after)}$ = roughness after grading, in QI;

$QI_{(before)}$ = roughness before grading, in QI;

$a = 0.553 + 0.230\, MGD_j$;

$QIMIN_j = \max\{10; \min[100; 4.69\, D95_j\,(1 - 2.78\, MG_j]\}$;

MG_j = slope of mean material gradation, such that

$MG_j = \min(MGM_j, 1 - MGM_j, 0.36)$, where

$$MGM_j = \frac{MG075_j + MG425_j + MG02_j}{3}, \quad \text{where}$$

$$MG075_j = \begin{cases} \ln(P075_j/95)/\ln(0.075/D95_j) & \text{if } D95_j > 0.4, \\ 0.3 & \text{otherwise;} \end{cases}$$

$$MG425_j = \begin{cases} \ln(P425_j/95)/\ln(0.425/D95_j) & \text{if } D95_j > 1.0, \\ 0.3 & \text{otherwise; and} \end{cases}$$

$$MG02_j = \begin{cases} \ln(P02_j/95)/\ln(2.0/D95_j) & \text{if } D95_j > 4.0, \\ MG425_j & \text{otherwise.} \end{cases}$$

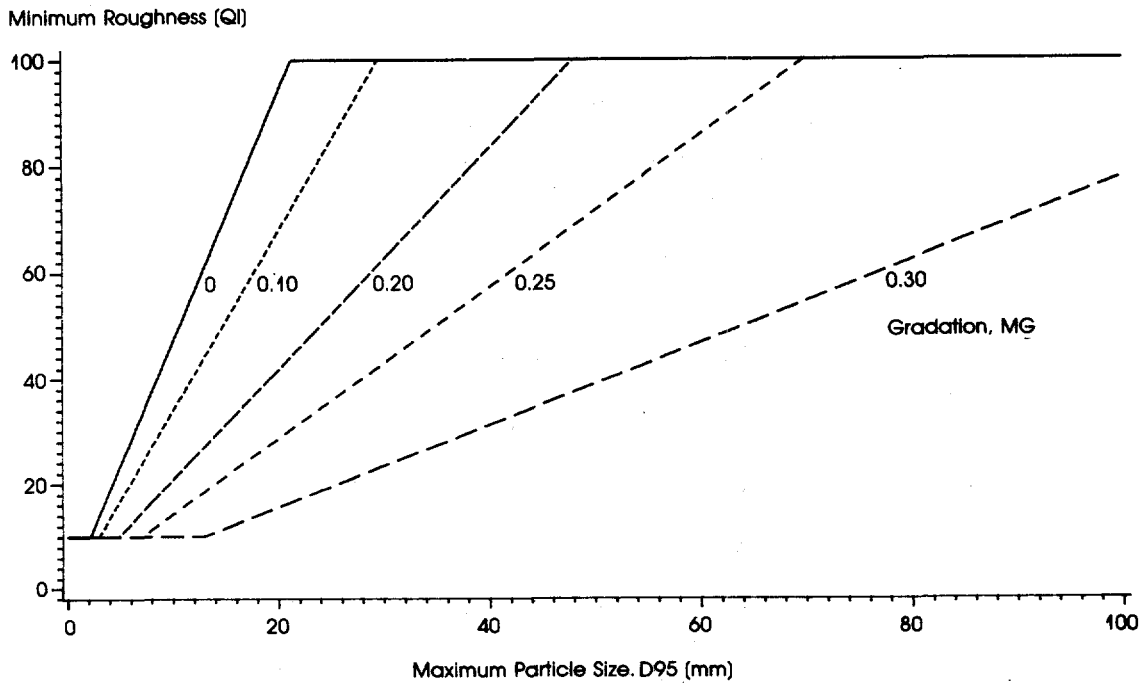
Note: The standard error of this prediction on the original data base was 31.6 QI (2.4 m/km IRI).

Average roughness in analysis year

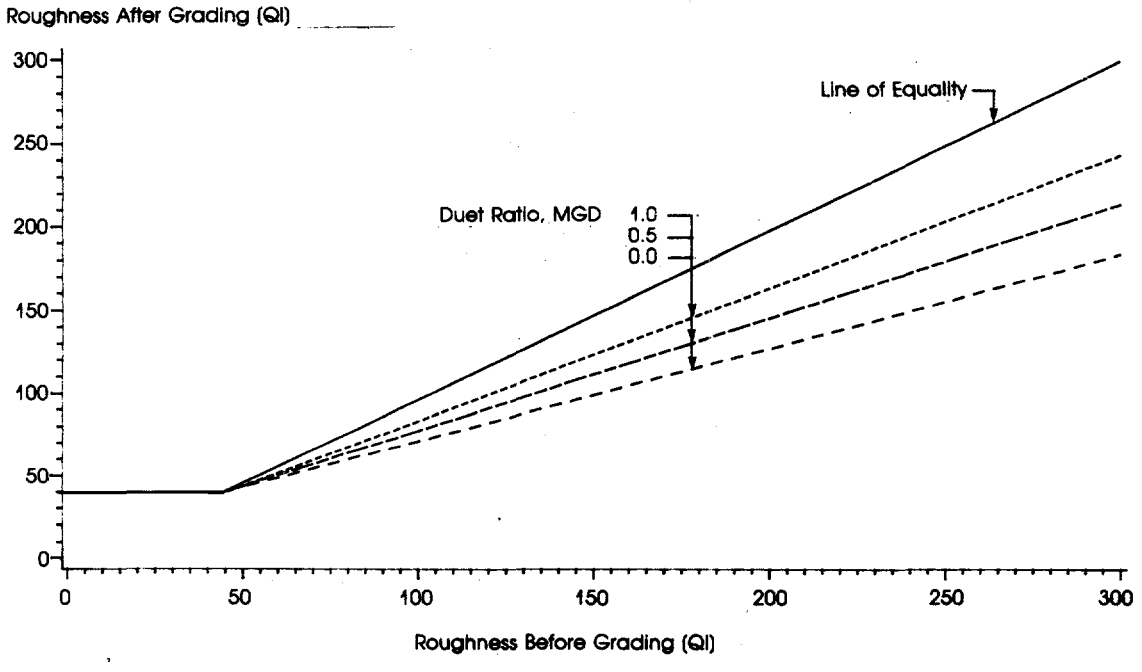
The average roughness during the analysis year is computed by combining the progression and grading-effect relationships and integrating (see Paterson 1987). The year's average is expressed in terms of the roughness at the beginning of the year and the parameters in the previous expressions as follows:

Figure 4.17: Prediction of minimum roughness and roughness after grading for unpaved roads

(a) Estimated minimum roughness for various materials



(b) Roughness after grading as a function of roughness before grading



Source: This study.

Case 1 ($t\ n \geq 1$)

The average roughness during year t, QI_{avg}, is given by:

$$QI_{avg} = QIMAX\ (1-y) + S_N\ y/n$$

where

$$y = (b-1)\ n/365\ c$$

$$S_N = [n\ k + (1-(a\ b)_n)\ QI_a - k\ (1-(a\ b)_n)/(1-a\ b)]/(1-a\ b).$$

$$k = (1-a)\ QIMIN + a\ (1-b)\ QIMAX$$

QI_a = roughness at beginning of year t where

$$QI_a = \begin{cases} QI_0 = \text{value specified by user or by default, in QI; if } t = 1; \\ \text{or} \\ QI_b = \text{roughness at end of year, } t - 1\text{, as given below in QI} \end{cases}$$

QI_{avg} = average roughness during year t, in QI

a, b, c = as defined above, except that b, c take the values b' and c' when mechanical compaction has been specified and the roughness at the end of the year, QI_b, is

$$QI_b = (ab)^n QI_a + k\ (1-(ab)^n)/(1-ab)$$

Case 2 ($t\ n < 1$)

$$QI_{avg} = QIMAX - (QIMAX-QI_a)\ [\exp(365\ c) - 1]/(365\ c)$$
$$QI_b = QIMAX - (QIMAX - QI_a)\ \exp(365\ c)$$

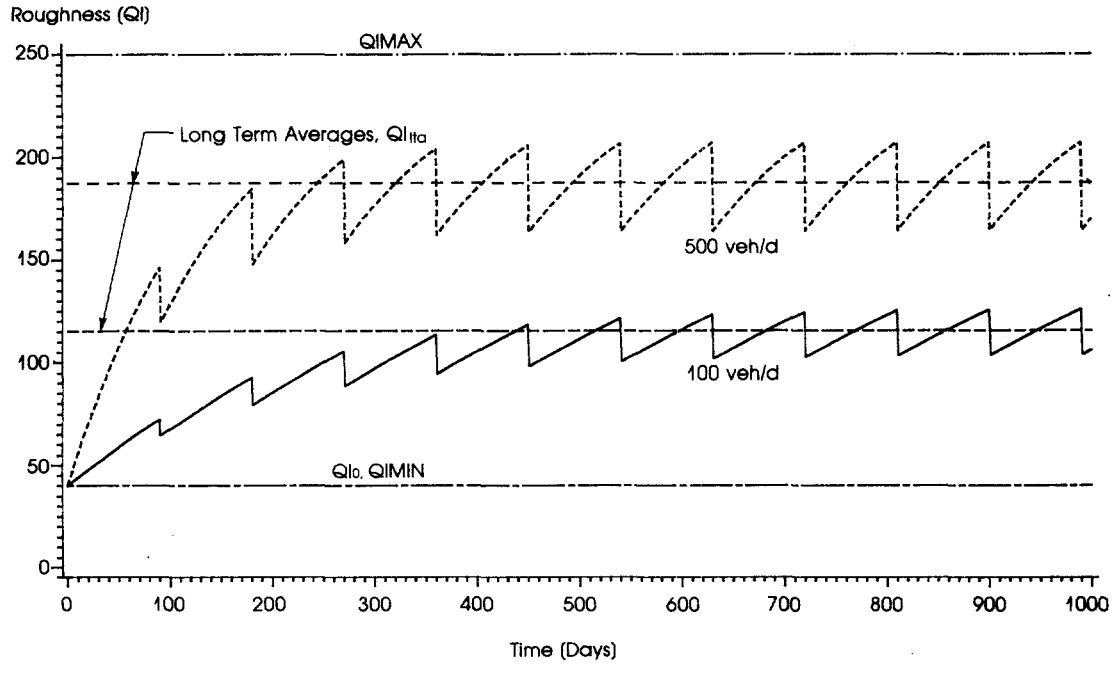
Roughness cycle "steady state"

When grading is performed regularly at constant time intervals, or a fixed roughness level, or fixed traffic intervals, the process of roughness change described by these relationships without restriction eventually leads to a steady state, as shown in Paterson (1987). This steady state is characterized by a saw-toothed pattern of roughness-time profile, in which the highs and lows represent the roughness immediately before and after grading, respectively. These highs and lows, denoted by QIH and QIL, are given by:

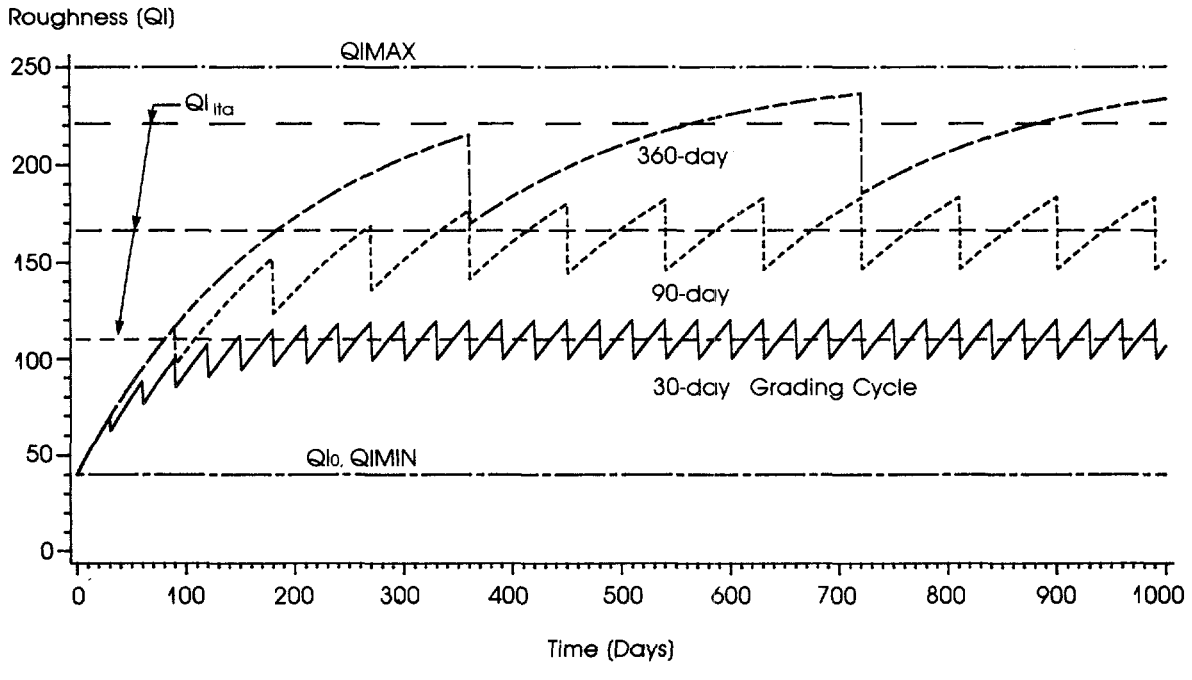
$$QIH = [QIMAX_j\ (1 - b) + QIMIN_j\ (1 - a)\ b]/(1 - a\ b);$$
$$QIL = [QIMIN_j\ (1 - a) + QIMAX_j\ a\ (1 - b)]/(1 - a\ b);$$

Figure 4.18: Predictions of roughness progression under various traffic volumes and grading policies for unpaved roads

(a) Effects of traffic volume under regular 90-day grading policy



(b) Various grading policies for average daily traffic 300 veh/day



Source: After Paterson (1987).

where all parameters are as defined above.

The long-term average roughness, denoted by QI_{lta}, at this steady state under a maintenance policy is dependent on the grading frequency (embodied in the parameter b above) and is obtained by integration over the roughness-time profile, so the annual average roughness tends to the following:

$$QI_{avg} \rightarrow QI_{lta}$$

where $$QI_{lta} = QIMAX_j + (1 - a)(1 - b)[QIMAX_j - QIMIN_j]/[(1 - a\,b)\,\ln b].$$

The relationships are illustrated in Figure 4.18 where the long-term average roughness is superimposed on the cyclic trends for (a) a road under regular 90-day grading maintenance and different levels of traffic, and (b) a road under different (30, 90, 360-day) grading policies and one level of traffic (300 veh/day : 200 light, 100 heavy). The surfacing material is a medium (20 mm) slightly plastic gravel with high dust ratio (0.80) and moderate gradation (MG = 0.20).

4.5.2 Material Loss

From the Brazil-UNDP study the following relationship for predicting the annual quantity of material loss as a function of monthly rainfall, traffic volume, road geometry and characteristics of the gravel (if gravel road) and the subgrade (if earth road) was obtained (Paterson, 1985):

$$MLA = 3.65\,[3.46 + 0.246\,MMP\,RF + KT\,ADT]$$

where MLA = the predicted annual material loss, in mm/year;

KT = the traffic-induced material whip-off coefficient, expressed as a function of rainfall, road geometry and material characteristics:

$$KT = \max\left[0;\ \left(0.022 + 0.969\,\frac{C}{57300} + 0.00342\,MMP\,P075_j - 0.0092\,MMP\,PI_j - 0.101\,MMP\right)\right];\ \text{and}$$

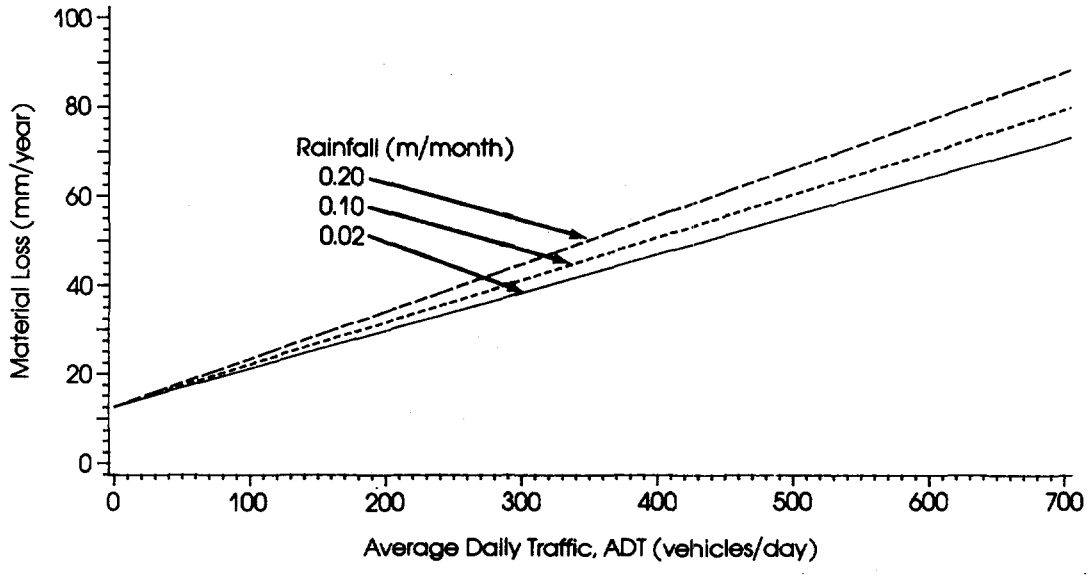
j = g if gravel road; ≃ s if earth road.

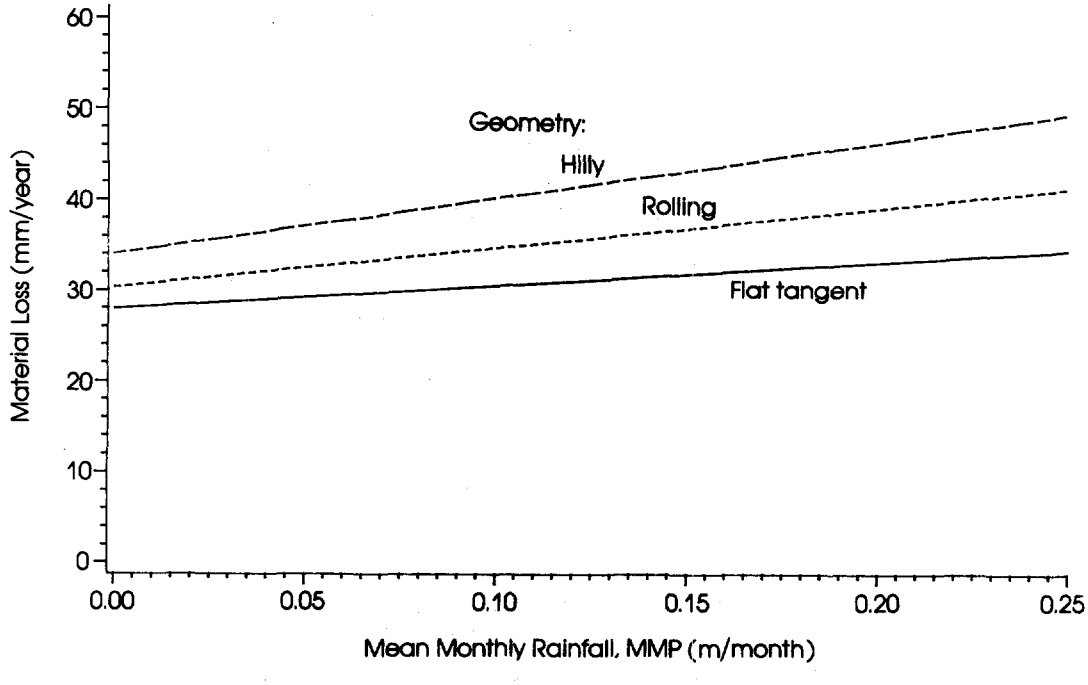
The predictions are illustrated in Figure 4.19 showing the effects of (a) traffic and rainfall for flat terrain and (b) rainfall and geometry for a traffic volume of 200 veh/day, and slightly plastic, fine silty gravel surfacing material.

4.5.3 Passability

Passability is the quality of the road surface which ensures the safe passage of vehicles. In the vehicle operating cost submodel, provision has been made to determine the economic impact of a partial

Figure 4.19: Predictions of surfacing material loss related to traffic, rainfall and geometry for unpaved roads





reduction in passability through factors augmenting the operating costs of the various vehicle types (see Section 5.3.7). This augmentation comes into effect when the gravel surfacing thickness drops below a minimum, and relates to the risk of the subgrade material being impassable.

The user however must determine exogenously whether passability will be a problem in the subgrade material, because no physical estimation of it is made within the model. The following criteria from Visser (1981) are adequate for ensuring passability and surface stability:

1. Passability which is a function of the shear strength of the saturated material, is satisfactory when:

   $SFCBR \geq 8.25 + 3.75 \log_{10} (ADT)$, and

2. Surfacing stability, which relates to ravelling and looseness, is satisfactory when:

   $P075 \geq 14$,

   where SFCBR = the (minimum) soaked California Bearing Ratio at standard Proctor laboratory compaction for ensuring passability; and

   ADT, P075 = as defined previously.

4.5.4 Grading Maintenance Options

The principal routine maintenance for unpaved roads, other than the miscellaneous items covered in Section 4.5.7, is grading which may be specified by the user in one of three ways, that is:

1. Scheduled: a fixed time interval in days between successive gradings;

2. Traffic-responsive: a fixed traffic interval in number of vehicle passes between successive gradings; and

3. Roughness-responsive: a maximum allowable roughness.

In all cases, the average roughness between successive gradings, QIAVG, is computed as a function of the number of days between gradings, DG, as described in Section 4.5.1. In the scheduled case, DG is specified directly by the user. In the traffic and roughness-responsive cases, DG is determined as follows:

$$DG = \begin{cases} DGMAX & \text{if } DGMAX < DG' \\ DG' & \text{if } DGMIN < DG' \leq DGMAX \\ DGMIN & \text{if } DG' \leq DGMIN \end{cases}$$

where DGMAX = the maximum allowable time interval between successive gradings, in days, specified by the user as an option or equal to the default value of 10,000;

DGMIN = the minimum applicable time interval between successive gradings, in days, specified by the user as an option or equal to the default value of 5 days;

DG' = the number of days between successive gradings determined from the traffic or roughness parameter, as follows:

$$DG' = \begin{cases} \dfrac{VEHG}{ADT} & \text{for the traffic-responsive case} \\ (1/c)\ \ell n\ \{(QIMAX_j - QIMAX_0)/[QIMAX_j - (1 - a)\ QIMIN_j - a\ QIMAX_0]\} & \text{for the roughness-responsive case} \end{cases}$$

VEHG = the traffic interval between successive gradings, in vehicles, specified by the user; and

$QIMAX_0$ = the maximum allowable roughness specified by the user, in QI.

No grading

If no grading is specified, the long-term average roughness is equal to the maximum roughness, as follows:

$$QI_{1ta} = QIMAX_j$$

Note that if the historic maintenance of the link has been nil-grading over several years, then the existing roughness is the best estimate of the average roughness and the user can provide this by specifying QIMAX exogenously with a value equal to the existing roughness.

4.5.5 Spot Regravelling

Spot regravelling provides repair to areas of severe depression (gravel loss, rutting, etc.), and may be specified by the user either in a fixed number of cubic meters per kilometer per year or as a percentage of gravel or subgrade material loss in the current analysis year to be replaced (subject to a maximum limit per year). When spot regravelling is performed, the added material is assumed to be the same type as the existing. The cost of spot regravelling is computed as the product of the unit cost of material per m^3 and the volume of material added in m^3 per km. For gravel roads the thickness of the gravel layer is increased to reflect the volume of material added, according to the following formula (trapezoidal rule):

$$\Delta THGS = \frac{VGS}{W + WS}$$

where $\Delta THGS$ = the increase in gravel thickness due to spot regravelling, in mm; and

VGS = the in-place volume of gravel added due to the spot regravelling, in m^3/km.

The spot regravelling is predicted to reduce the average roughness on the assumption that the gravel is applied in the major depressions and potholes that have appeared in the surface in the upper ranges of roughness. Roughness levels above 190 QI (15 m/km IRI) are invariably associated with the presence of visible birdbath type depressions or potholes, which become larger or more frequent as the roughness level increases, and these can be effectively patched, with high benefits, by spot regravelling. Over the roughness range of 150 to 190 QI (11 to 15 m/km), such "patchable" birdbath depressions are frequently but not always present so that, in this range, spot regravelling may not always be effective. For example, spot regravelling is not effective maintenance on corrugations or on runoff-induced surface erosion, which conditions commonly induce roughness levels within this range. At roughness levels below 150 QI (11 m/km IRI) spot regravelling is considered to be ineffective on roughness. This logic is defined in the following algorithm, adopting the same roughness : volume of depression ratio as for paved road pothole simulation (i.e., 2 QI per m^3/lane/km), allowing for the spot regravelling to be only 60 percent effective (i.e., 1.2 QI per m^3/lane/km), and adopting an average effective "lane" width of 3 m:

$$QI_{avg(after)} = \max\{150;\ [QI_{avg(before)} - \min(1;\ (QI_{avg(before)} - 150)/40)\ VGS\ 3.6/W]\}.$$

This is illustrated in Figure 4.20. It should be noted that spot regravelling affords only a temporary repair of depressions, and that the most effective means is by grading, or in severe cases by scarifying, grading and recompacting.

4.5.6 Gravel Resurfacing Maintenance

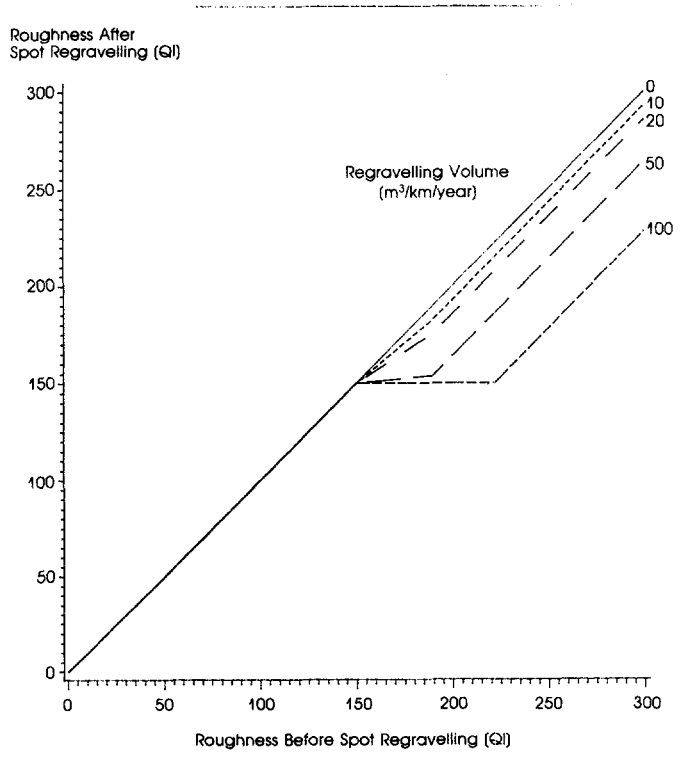
A gravel resurfacing policy is specified either as:

1. Scheduled, in which gravel resurfacing is applied when the gravel age, GAGE, equals or exceeds the fixed time interval specified by the user; or

2. Condition-responsive, in which gravel resurfacing is applied when either

   (i) The current gravel thickness, THG, falls below the user-specified minimum allowable thickness, provided that the gravel age, GAGE, equals or exceeds the (optional) user-specified minimum applicable resurfacing interval, in years; or

Figure 4.20: Effect of spot regravelling on average roughness



Source: This study.

(ii) The gravel age, GAGE, equals or exceeds the (optional) user-specified maximum allowable resurfacing interval, in years.

However, gravel resurfacing is not performed if the specified last applicable year has been exceeded, or if the analysis year is a construction year, or if the final thickness specified in a scheduled policy is smaller than the predicted thickness at the end of the analysis year before the resurfacing decision.

When gravel resurfacing is performed the surface type is set to "gravel" (regardless of the previous surface type), the gravel age, GAGE, is reset to zero, the existing gravel material is changed to the material specified by the user (which may still be of the same attributes as the existing) and the thickness of the gravel surface is increased according to the formula below:

$$THG_{(after)} = \begin{cases} THG_0 & \text{if the final gravel thickness is specified} \\ THG_{(before)} + \Delta THG_0 & \text{if an increase in the gravel thickness is specified} \end{cases}$$

where THG_0 = the user-specified final gravel thickness after resurfacing, in mm; and

ΔTHG_0 = the user-specified increase in the gravel thickness due to resurfacing, in mm.

The volume of gravel added per km is computed according to the following trapezoidal formula:

$$VGR = [THG_{(after)} - THG_{(before)}] (W + WS)$$

where VGR = the in-place volume of gravel material added due to gravel resurfacing, in m^3/km.

Finally, the values of the gravel attributes ($PO75_g$, $P425_g$, $PO2_g$, $D95_g$, PI_g, $QIMIN_g$ and $QIMAX_g$), are replaced either by the new values provided by the user, or by default values from the previous gravel attributes.

4.5.7 Routine-Miscellaneous Maintenance

This includes drainage maintenance, vegetation control, shoulder maintenance, safety installations, and other items which are not modelled as affecting the riding quality of the pavement. A lump sum cost per km per year is used as the basis for costing routine maintenance. Because the unpaved road deterioration relationships employed are based on the assumption of adequate drainage, the cost of drainage maintenance should be included, when it is normally done; otherwise, some allowance due to the lack of drainage, e.g., in the form of frequent road closures, washouts, etc., should be incorporated in the economic analysis.

CHAPTER 5

Vehicle Operating Cost Submodel

5.1 GENERAL OUTLINE

5.1.1. Operation of the Submodel

The function of the vehicle operating cost submodel is to simulate the effects of the physical characteristics and condition of a road on the operating speeds of various types of vehicles, on their consumption of fuel and lubricants, on their maintenance requirements, and so on, and to determine their total operating costs. The quantities of resources consumed, such as liters of fuel, numbers of tires, man-hours of labor, etc., are determined together with vehicle speeds as functions of the characteristics of each type of vehicle and the geometry, surface type, and current condition of the road. Costs are then found by multiplying the various resource quantities by user-specified unit costs and adding allowances for depreciation, interest, and overhead costs and for the time values of passenger delays and cargo holding.

The user may specify prices or unit costs in both financial and economic terms and may specify the foreign exchange element of total costs. Financial costs represent the actual costs incurred by transport operators in owning and operating vehicles over the road. Economic costs represent the real costs to the economy of that ownership and operation, where adjustments are made to allow for market price distortions such as taxes, foreign exchange restrictions, labor wage laws, etc., and where the implicit costs of passengers' time and cargo holding are accounted for. Foreign exchange costs represent the costs which must be provided for in foreign currencies and, depending on the use, can be interpreted as either financial or economic.

The procedures followed by the vehicle operating cost submodel in computing speeds, resource use, and costs for the traffic using a given road section in each year may be summarized as follows:

1. Compute the average operating speed for each vehicle group.

2. Compute the amounts of resources used per vehicle-kilometer by each group for the following components:
   a. Fuel
   b. Tire wear
   c. Maintenance parts
   d. Maintenance labor
   e. Lubricants
   f. Crew
   g. Depreciation
   h. Interest

i. Overhead
j. Passenger time
k. Cargo holding
l. Miscellaneous costs

3. Apply unit costs to the resource consumption amounts to obtain cost per vehicle-kilometer for each vehicle group.

4. Multiply cost per vehicle-kilometer by the section length and by the year's traffic volume of each vehicle group to obtain the total road user costs for the year for each group.

5. Sum the group totals to obtain the overall total road user costs for the year.

Throughout any computer run, the values of the parameters that describe the environment, vehicle characteristics, and unit costs remain unchanged, while the variables that describe traffic flows, traffic composition, road geometry, and the type and condition of the surface may change from year to year according to the actions of the submodels described in previous chapters.

In the HDM-III model there are actually four different sets of relationships for estimating vehicle speeds and the consumption of some of the operating resources, based on four separate empirical studies by different agencies. The user must choose and specify which of these sets is to be employed in any given run. We give guidance on that issue in the next section. We then describe the relationships derived from the study conducted in Brazil. Also in this chapter, following the Brazil relationships for certain components, are the relationships that are not differentiated by the different studies or data sources and which are used regardless of which of the four options is chosen. Relationships developed by British Transport and Road Research Laboratory (TRRL) from studies in Kenya and the Caribbean, and by the Central Road Research Institute of New Delhi, (CRRI) from research in India, are presented in Chapter 6.

5.1.2 Choice of Relationships

The principal set of relationships are those derived from the Brazil study by GEIPOT, the Texas Research and Development Foundation, and the World Bank. The alternative relationships are those from the Kenya and Caribbean studies by the British Transport and Road Research Laboratory, and those from the India study by the Central Road Research Institute-New Delhi. For brevity, the relationships derived from the Brazil studies are referred to as the Brazil relationships, and the alternative sets of relationships are referred to as the Kenya, Caribbean, and India relationships.

Chapter 1 has already provided some discussion of the various sets of relationships, and certain further salient features are discussed here. For a more comprehensive discussion of the four studies, including a comparison of the resulting relationships, the reader is referred to Chesher and Harrison (1987).

Appendix 5A provides comparative information for the four studies concerning coverage of vehicle types (Table 5A.1) and the required information for utilizing the different relationships with respect to vehicle characteristics (Table 5A.2) and road characteristics (Table 5A.3). By examining these tables one can gain some further appreciation of the differences in the various sets of relationships.

Vehicle types. The vehicle types covered by the different studies differed significantly. In the Caribbean study, only four categories were distinguished, and there were no heavy or even medium trucks or buses. At the other extreme, the Brazil study used ten categories, including three classes of cars and five of trucks. Table 5A.1 lists the characteristics of the different vehicle types included in each of the four studies.

Model formulation. The studies also differed considerably in the degree of detail in the data and in the methods of analysis and formulation used. The relations using the most detailed sets of explanatory variables are those from Brazil. In that study the formulation of relations for vehicle speed, fuel consumption, and truck and bus tire wear were based on generally accepted principles of vehicle mechanics and driver behavior. With that theoretical basis and the explicit accounting for many explanatory variables, the relations should be applicable in principle to many situations where combinations of conditions differ from those observed in the field studies in Brazil, although it must always be recognized that differences in economic circumstances, not all of which are encapsulated in the models, can affect the relationships.

The other studies used fewer explanatory variables and relied more on statistical correlation of associated variables through linear multiple regression. This yielded simpler relations while implicitly assuming that some of the details accounted for in the Brazil formulation could be safely ignored, either as being insignificant or as being highly correlated with, and therefore accounted for by, variations in the explicit variables.

The different input data requirements and options of the different sets of relationships are indicated in Tables 5A.2 and 5A.3. Compared with the other three sets, the Brazil relations use many more explanatory variables. Wherever possible the user should obtain local information on these variables, as it will better calibrate the model to local conditions. For many of them, however, default values based on the empirical study will be automatically supplied if inputs are not entered, as indicated in the tables by the letter "O" for "optional." For these items the user does not have to provide inputs where there is not better local information or other basis for using different values; in such case, however, the user should be aware that he is employing potentially inaccurate assumptions. Since these parameters are already implicitly assumed in the Kenya, Caribbean and India relationships, the user has no alternative but to employ the same assumptions used in the original study regardless of whether better local information is available -- and similar cautions apply *a fortiori*.

Other considerations. In considering all the various factors -- physical, economic, and behavioral -- which may affect vehicle operating cost relationships, it is clear that the circumstances of India make it somewhat unique among the four countries studied. Vehicle designs are of older vintage with lower horsepower/weight ratios. Road and traffic flow conditions are much different, with generally poorer road characteristics and significant traffic-interference even under nominally free-flowing conditions, and Indian drivers have adapted behavior accordingly. Even more striking is the contrast in general economic conditions, particularly the relative costs of capital and labor and restrictions in the availability of new vehicles. These different economic circumstances undoubtedly constrain, and probably dominate, decisions concerning vehicle maintenance and replacement. Circumstances force vehicle owners to maintain older vehicles (with much higher expenditures on labor and lower expenditures on spare parts due to the cheapness of labor) for much longer periods than they do in the other countries. Since for vehicle maintenance costs we do not yet have adequate theoretical models, and are forced to rely on correlation rather than mechanistic-behavioral type models for this component, particular care should be taken in local calibration of this component, and under Indian type conditions the India relationships may be preferred. Further research is planned to compare predictive abilities of the Brazil models calibrated for Indian conditions against the CRRI India models.

Conclusions on the choice of vehicle operating costs relationships. As indicated in Chapter 1, because of their more general form and more extensive empirical validation, the user is advised normally to use the Brazil relationships with as much local calibration as possible. (See Chapter 13, Watanatada et al., 1987, for guidance on local calibration.)

The principal exceptions would be for applications in the Caribbean, Kenya and particularly India, where the alternative relationships were themselves statistically estimated. In these cases it is possible that local variations in economic circumstances, traffic conditions and driver behavior may be better captured in the relationships estimated from the substantial local research studies than by the more general model forms drawn from Brazil and locally calibrated.

It should always be borne in mind, however, that the ultimate concern is to predict accurately the changes (or first differences) in vehicle operating costs with respect to changes in road characteristics, rather than total operating costs over averaged conditions. Cost differentials due to different road conditions are expected to vary much less than total vehicle operating costs in response to different economic environments (Chesher and Harrison, 1987).

At the present stage of the research, we are unable to give more specific guidance, except to recommend that an HDM model user in India, Kenya or the Caribbean should apply the locally calibrated Brazil models as well as the locally established relationships and compare the results to gain a broader insight into probable effects of different road characteristics.

5.2 THE BRAZIL RELATIONSHIPS

The Brazil portion of the vehicle operating cost (v.o.c.) submodel is made up of relationships based on the Brazil-UNDP study (GEIPOT, 1982; Chesher and Harrison, 1985; Watanatada et al., 1985). The vehicles observed in this study were classified in the ten groups shown in Table 5A.1 (in Appendix 5A), where the representative characteristics of each vehicle type are tabulated. The forms of the relations for predicting vehicle speed, fuel consumption, and bus and truck tire wear are based on principles of vehicle mechanics and driver behavior (Watanatada, et al., 1987), while those for predicting maintenance parts and labor requirements are based on econometric analysis of user survey data (Chesher and Harrison, 1987). These relations are explained in the following sections.

5.2.1 Vehicle Speeds

The prediction of vehicle speeds is based on what may be described as an aggregate probabilistic limiting velocity approach to steady-state speed prediction. The term aggregate connotes that the prediction method works with aggregate descriptors of road geometry and surface condition rather than with detailed information about the roadway. The expression steady-state implies that transitional effects, that is, speed-change cycles along the roadway, are not modeled. The approach is termed a probabilistic limiting velocity approach because the predicted speed is a probabilistic minimum of a number of limiting or constraining speeds. These constraining speeds are functions of such factors as characteristics of the vehicle (e.g., engine power, braking capacity) and of the roadway (e.g., vertical gradient, roughness). Here we give a brief summary statement of the model; for a more detailed presentation of the methodology and its validation the reader is referred to Chapters 3 through 8 of Watanatada et al. (1987).

For the purpose of predicting the average vehicle speed for a round trip, it has proved satisfactory to represent the given roadway of various vertical and horizontal alignments by two idealized homogeneous segments -- one of positive grade (uphill) and the other of negative grade (downhill), both having the same length, roughness, average horizontal curvature, and average rise plus fall.

The steady-state speed for each type of vehicle is predicted for each of these hypothetical road segments as a function of the road characteristics and the attributes of the vehicles. Then the average round trip speed is computed to correspond to the space-mean speed over the two segments (i.e., the round trip distance divided by the round trip time).

The information on the roadway needed for the application of the speed prediction method is: the surface type of the roadway, the average roughness, an aggregate measure of the vertical gradient of the roadway, an aggregate measure of its horizontal curvature and an average value of superelevation. The roughness value is the average of values measured over short homogeneous subsections of the actual roadway. In HDM-III this value is internally computed as described in Chapter 4. The aggregate measures of vertical and horizontal alignment are defined as follows:

The gradient is expressed as the average rise plus fall, RF, which is defined as the sum of the absolute values, in meters, of all ascents and all descents along a section, divided by the length of the section in kilometers. (This is numerically equal to ten times the mean absolute gradient in percent.) The concept is illustrated in Figure 5.1.

The average horizontal curvature, C, is defined as the sum of the absolute values of angular deviations (in degrees) of successive tangent lines of the road alignment when traveling in one direction, divided by the section length in km. This concept is also illustrated in Figure 5.1. Note that the denominator is the arc length (not the chord length) of the section.

The average superelevation, SP, is defined as the weighted average of superelevations (in percent) of the curvy sections of the roadway, the weights being the proportions of the lengths of the curvy sections. Note that the vertical and horizontal measures are independent of direction along the road section.

The prediction of a vehicle's steady-state speed on a given road segment makes use of a set of limiting (or "constraining") velocities, corresponding to several different factors that tend to limit the speed. We first define the variables, then give a diagrammatic discussion, and, finally, a formal mathematical statement of the model.

Variables

The constraining velocities, which are explained in subsequent sections, are (all in meters/sec):

VDRIVE = the limiting speed based on vertical gradient and engine power,

VBRAKE = the limiting speed based on vertical gradient and braking capacity,

VCURVE = the limiting speed determined by road curvature,

VROUGH = the limiting speed based on road roughness and associated ride severity, and

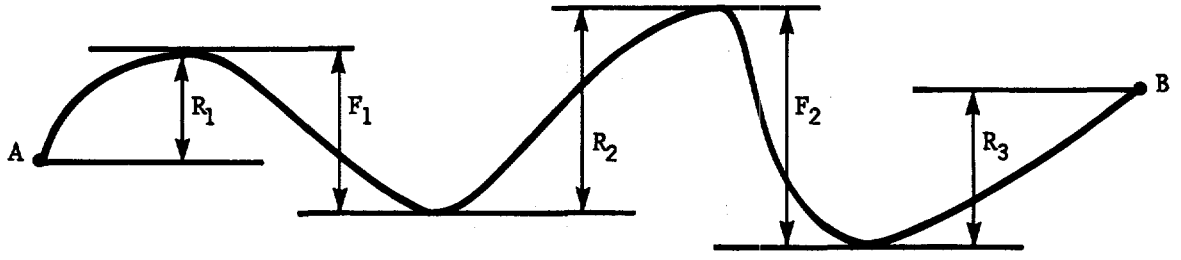
VDESIR = the desired speed in the absence of other constraints, based on psychological, economic, safety, and other considerations.

The two gravity-related constraining speeds, viz. VDRIVE and VBRAKE, have different values for the uphill and the downhill segments representing a given roadway. Each of the other three constraining speeds has the same value for both segments.

Diagrammatic exposition. The three plots in Figure 5.2 illustrate the constraining speeds and the resulting steady-state speed on

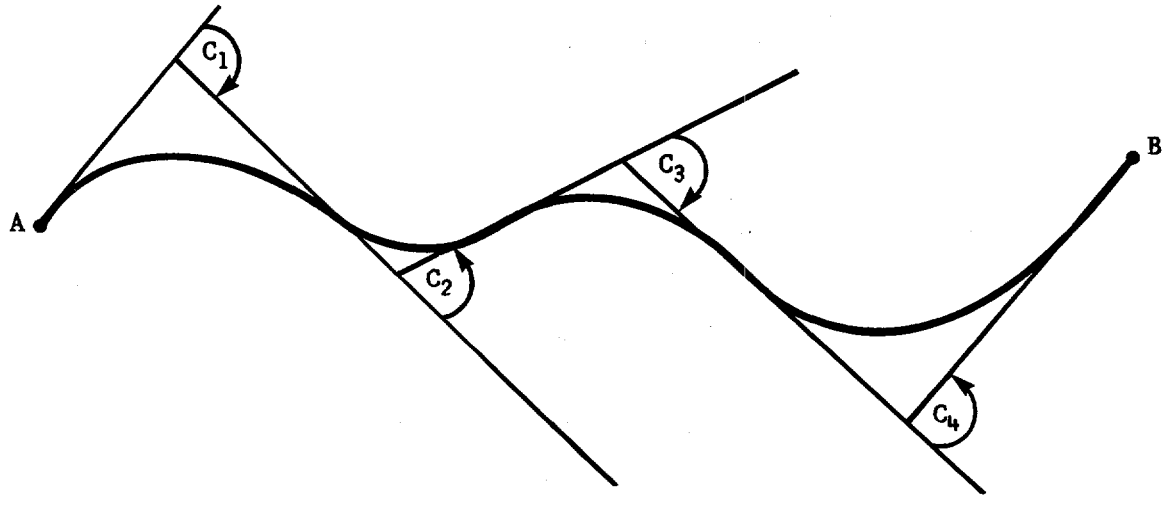
Figure 5.1: Illustration of road rise plus fall and horizontal curvature

(a) Vertical profile of the road section



$$\text{Average rise plus fall, RF} = \frac{R_1 + R_2 + R_3 + F_1 + F_2}{L_{ab}} \quad \frac{\text{(meters)}}{\text{(km)}}$$

(b) Plan view of the road section



$$\text{Average horizontal curvature, C} = \frac{C_1 + C_2 + C_3 + C_4}{L_{ab}} \quad \frac{\text{(degrees)}}{\text{(km)}}$$

a segment for a heavy truck carrying a net load of 6,000 kg, and traversing a paved segment. Each of the plots shows the response of the steady-state speed as one of the three speed-influencing factors is varied holding the other two factors constant.

Figure 5.2(a) shows the effect of surface irregularity on the steady-state speed for a straight and downward-sloping segment. The steady-state speed ($\hat{V}$) and the constraining speed based on road roughness (VROUGH) are graphed using continuous lines and the other limiting speeds using broken lines. At any given point on the roughness axis, the limiting speed with the lowest value may be considered the "binding" speed as it exercises the predominant influence on the resulting steady-state speed at that point; other limiting speeds have only marginal effects. The higher the value of a limiting speed, the more marginal its influence on the predicted steady-state speed. In this instance, over the lower range of road roughness (QI under about 83), the desired speed (VDESIR) becomes the binding speed. It may also be noted that roughness has no influence on VDESIR and VCURVE, but has a slight influence on the gravity-related constraining speeds, VDRIVE and VBRAKE through the rolling resistance coefficient.

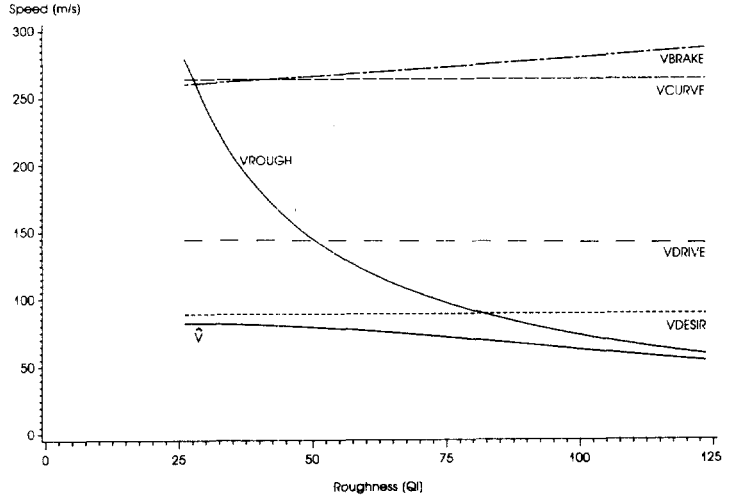
The effect of vertical alignment on the steady-state speed for slightly curved and low roughness segments is shown in Figure 5.2(b). The two constraining speeds most closely associated with road gradient, namely, VDRIVE and VBRAKE, as well as steady-state speed, $\hat{V}$, are shown using continuous lines. In this figure, three different constraining speeds are binding over the $\pm$10 percent range of the road gradient. On the one extreme, for negative grades steeper than 7.5 percent, the limiting speed based on braking capacity (VBRAKE) is binding. At the opposite end, the limiting speed based on engine power (VDRIVE) becomes dominant for slightly negative (-0.2 percent) and positive gradients. In the mid-range the steady-state speed is determined by the desired speed (VDESIR). It may be noted that for slightly negative and positive grades, the value of VBRAKE is infinity, that is, it has no influence on the resulting steady-state speed over this range, while road gradient has no influence on VDESIR, VCURVE, and VROUGH.

Figure 5.2(c) shows the effect of horizontal alignment on the steady-state speed for smooth and downward-sloping segments. The steady-state speed and the limiting speed determined by road curvature (VCURVE) are plotted using continuous lines. In this case two constraining speeds are binding. The desired speed is the primary determinant over gentle curvature (up to about 250 degrees/km), beyond which the curve speed constraint prevails. VDESIR, VBRAKE, VDRIVE and VROUGH are all independent of horizontal alignment.

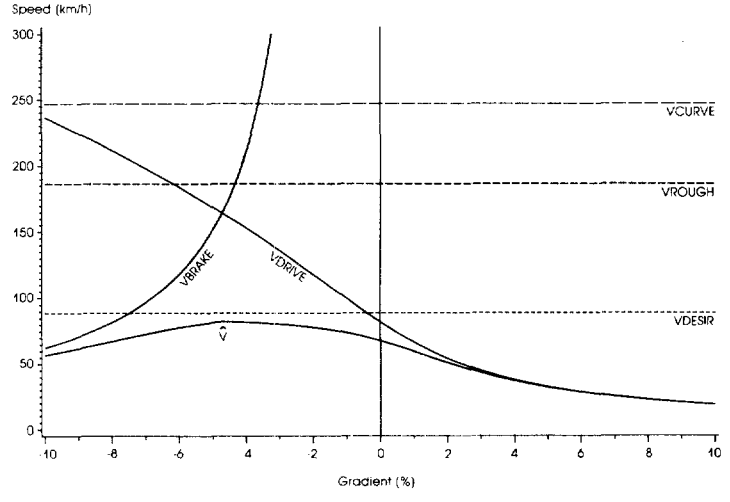
Mathematical summary. Using the respective values of the five limiting speeds for each road segment, the predicted steady-state speed for the segment is computed. The theory behind these computations involves treating each of the limiting speeds for a segment as a random variable and

Figure 5.2: Effect of constraining factors on predicted steady-state speed

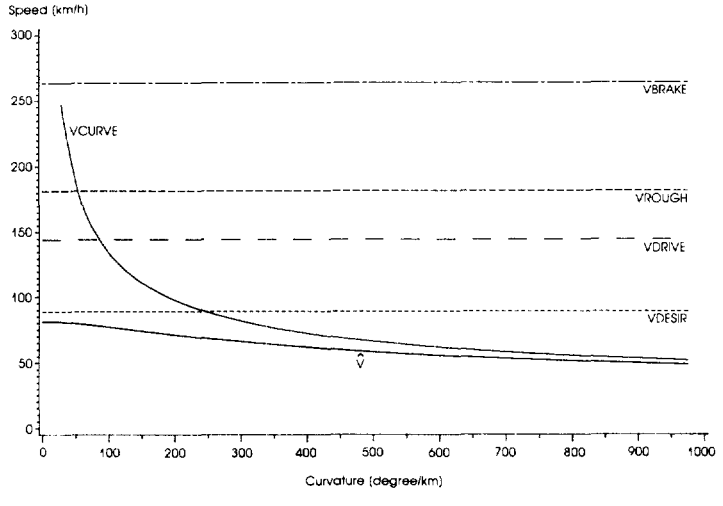
(a) Surface irregularity



(b) Vertical alignment



(c) Horizontal alignment



Source: Analysis of Brazil-UNDP-World Bank highway reseach project data (GEIPOT, 1982; Watanatada et al., 1987).

the resulting steady-state speed prediction as the average value of the minimum of these random variables. The probability model used is the Weibull distribution which is one of the standard "extreme value" distributions. The formulas are:

$$V_u = \frac{E^\circ}{\left\{\left(\frac{1}{VDRIVE_u}\right)^{\frac{1}{\beta}} + \left(\frac{1}{VBRAKE_u}\right)^{\frac{1}{\beta}} + \left(\frac{1}{VCURVE}\right)^{\frac{1}{\beta}} + \left(\frac{1}{VROUGH}\right)^{\frac{1}{\beta}} + \left(\frac{1}{VDESIR}\right)^{\frac{1}{\beta}}\right\}^{\beta}}$$

$$V_d = \frac{E^\circ}{\left\{\left(\frac{1}{VDRIVE_d}\right)^{\frac{1}{\beta}} + \left(\frac{1}{VBRAKE_d}\right)^{\frac{1}{\beta}} + \left(\frac{1}{VCURVE}\right)^{\frac{1}{\beta}} + \left(\frac{1}{VROUGH}\right)^{\frac{1}{\beta}} + \left(\frac{1}{VDESIR}\right)^{\frac{1}{\beta}}\right\}^{\beta}} .$$

In the above formulas, the subscripts u and d stand for the uphill and the downhill segment, respectively. It may be noted that only the two constraining velocities involving vertical gradient carry these subscripts. The coefficient β determines the shape of the assumed Weibull distribution and is a member of the set of parameters estimated for each type of vehicle. As the estimation involved a logarithmic transformation of the variables, the prediction formulas include the bias correction factor E°, equal to $\exp(0.5\ \sigma^2)$ where σ is an estimate of the standard error of residuals in the estimation.[1] The numerical values of β and E° as estimated from the Brazil data set and used in HDM are given in Table 5.1.

From the uphill and downhill speeds, V_u and V_d, the average speed for a round trip, using both segments, is calculated as follows:

$$S = \frac{(3.6)(2)\ LGTH}{\frac{LGTH}{V_u} + \frac{LGTH}{V_d}} = \frac{7.2}{\frac{1}{V_u} + \frac{1}{V_d}}$$

where LGTH = the length of the roadway in km;
V_u = the predicted vehicle speed for the uphill sedgment in m/s;
V_d = the predicted vehicle speed for the downhill segment in m/s;
3.6 = conversion factor from m/s to km/h.

---

1 An explanation of the correction of biases arising from non-linear transformation of variables in estimation is given in Chesher (1982).

Table 5.1: Parameter and default values for speed prediction

Parameters		Vehicle type								
		Small car	Medium car	Large car	Utility	Bus	Light truck		Medium/ heavy truck	Articulated truck
							gas	diesel		
Drag coefficient, CD		0.45	0.50	0.45	0.46	0.65	0.70	0.70	0.85	0.63
Frontal area, AR (m^2)		1.80	2.08	2.20	2.72	6.30	3.25	3.25	5.20	5.75
Payload, LOAD (tons)[1]		0	0	0	0.3	2.3	2.0	2.0	4.5/6.0[2]	13.0
HPDRIVE (metric hp)[3]		30.0	70.0	85.0	40.0	100.0	80.0	60.0	100.0	210.0
HPBRAKE (metric hp)[3]		17.0	21.0	27.0	30.0	160.0	100.0	100.0	250.0	500.0
FRATIO$_0$ (ton^{-1})[3]	Paved roads	0.268	0.268	0.268	0.221	0.233	0.253	0.253	0.292	0.170
	Unpaved roads	0.124	0.124	0.124	0.117	0.095	0.099	0.099	0.087	0.040
FRATIO$_1$ (ton^{-1})[3]	Paved roads	0	0	0	0	0	0.0128	0.0128	0.0094	0.0023
	Unpaved roads	0	0	0	0	0	0	0	0	0
ARVMAX (mm/s)[3]		259.7	259.7	259.7	239.7	212.8	194.0	194.0	177.7	130.9
VDESIR$_0$ (km/h)[3]	Paved roads	98.3	98.3	98.3	94.9	93.4	81.6	81.6	88.8	84.1
	Unpaved roads	82.2	82.2	82.2	76.3	69.4	71.9	71.9	72.1	49.6
B_W		0.74	0.74	0.74	0.74	0.78	0.73	0.73	0.73	0.73
β		0.274	0.274	0.274	0.306	0.273	0.304	0.304	0.310	0.244
E^*		1.003	1.003	1.003	1.004	1.012	1.008	1.008	1.013	1.018

[1] Excludes the weight of the two drivers (150 kg) which is already included in the tare weight.
[2] The payload is 4.5 tons for a medium truck and 6.0 tons for a heavy truck.
[3] These terms are defined in the following discussion in the text.

Source: Watanatada et al. (1987).

Note that the length term cancels out and the average predicted speed over the roadway is independent of its length.

Figures 5.3 and 5.4 show predicted average speed as a function of highway characteristics for a half-laden heavy truck over paved and unpaved roads, respectively.

VDRIVE

VDRIVE, the speed for a given road segment as limited by power and gradient, is arrived at from the hypothesis that the vehicle is driven at steady-state speed on a smooth, straight road employing a high level of power called the "used driving power", HPDRIVE. The used power has been found generally to be less than the rated power of the engine, especially for gasoline engined vehicles. Reasons for the difference are largely behavioral (unwillingness of drivers to use full power) and perhaps partly mechanical (operation at less than rated rpm, power lost in the transmission and used by accessories). HPDRIVE has been estimated from the Brazil speed observations, and the results for a number of vehicle types are included in Table 5.1.[2]

VDRIVE is related to HPDRIVE and the gradient through the balance of forces in the absence of acceleration:

$$\begin{bmatrix}\text{Drive}\\ \text{force}\end{bmatrix} = \begin{bmatrix}\text{Rolling}\\ \text{resistance}\end{bmatrix} + \begin{bmatrix}\text{Grade}\\ \text{resistance}\end{bmatrix} + \begin{bmatrix}\text{Air}\\ \text{resistance}\end{bmatrix}$$

where the various terms, all measured in newtons, are given by the following expressions:

$$\text{Drive force} = \frac{736 \quad \text{HPDRIVE}}{\text{VDRIVE}}$$

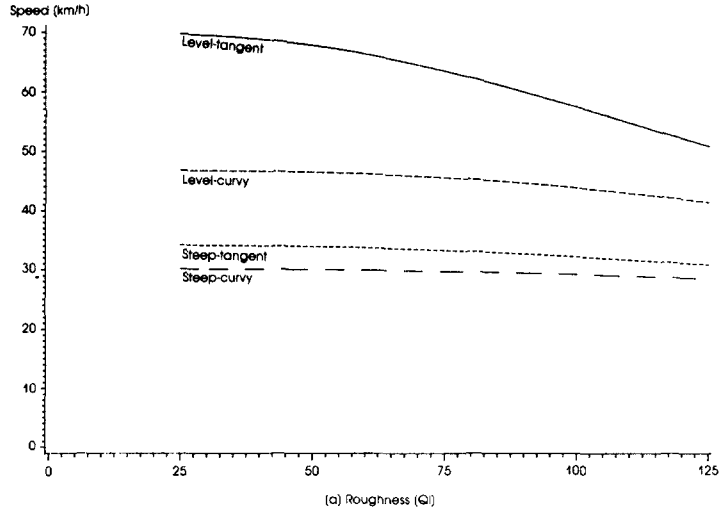
$$\text{Rolling resistance} = g\ 1000\ \text{GVW CR}$$
$$\text{Grade resistance} = g\ 1000\ \text{GVW GR}$$
$$\text{Air resistance} = 0.5\ \text{RHO CD AR VDRIVE}^2$$

where
- 736 = the number of watts in one metric hp;
- GVW = the gross vehicle weight, in tons;
- g = the gravitational constant, equal to 9.81 m/s^2;
- CR = the dimensionless coefficient of rolling resistance;
- GR = the vertical gradient expressed as a fraction;
- RHO = the mass density of air, in kg/m^3;
- CD = the dimensionless aerodynamic drag coefficient of the vehicle, which may be user-specified or take on a default value as shown in Table 5.1; and

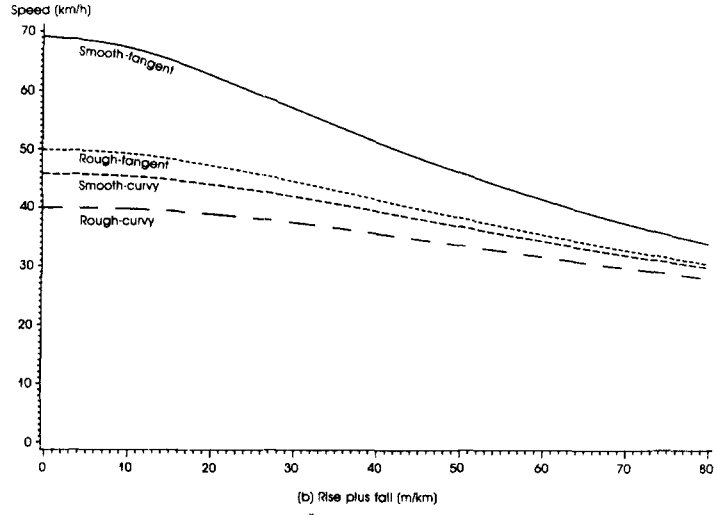
[2] Guidelines for calibrating HPDRIVE, as well as HPBRAKE and various other parameters, are given in Chapter 13 of Watanatada et al., 1987.

Figure 5.3: Speed as a function of road characteristics: half-laden heavy truck, paved road

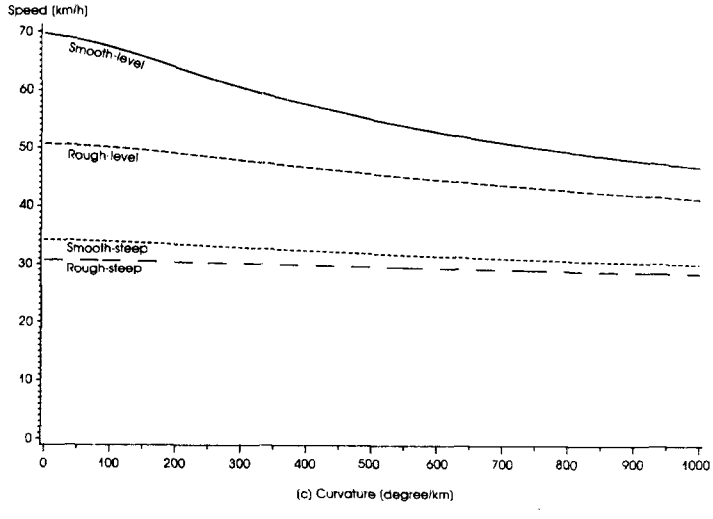
(a) Roughness



(b) Rise plus fall



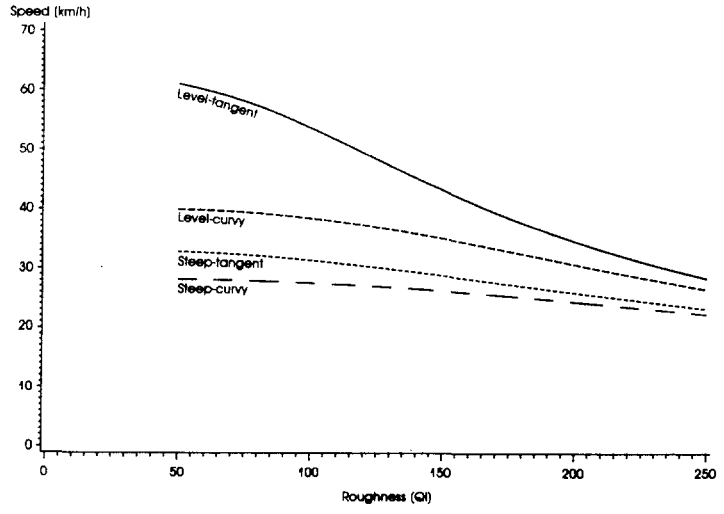
(c) Curvature



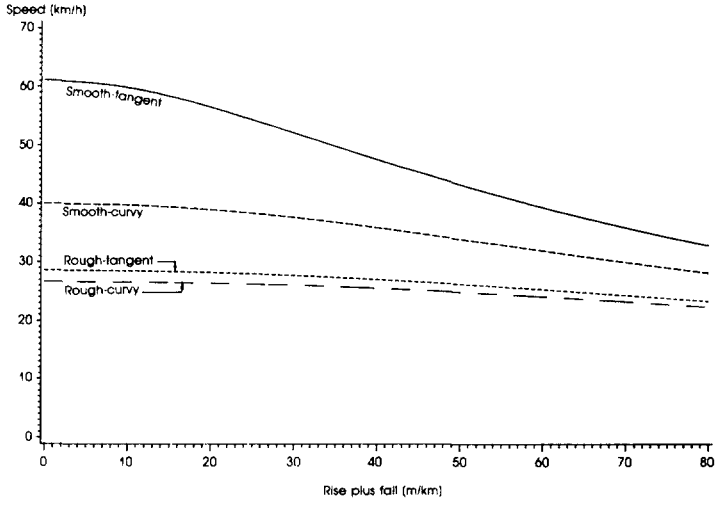
Source: Analysis of Brazil-UNDP-World Bank highway research project data (GEIPOT, 1982; Watanatada et al., 1987).

Figure 5.4: Speed as a function of road characteristics: half-laden heavy truck, unpaved road

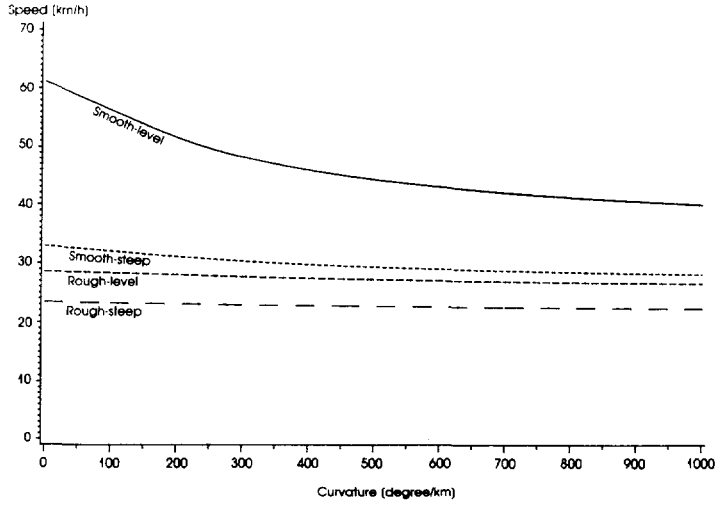
(a) Roughness



(b) Rise plus fall



(c) Curvature



Source: Analysis of Brazil-UNDP-World Bank highway research project data (GEIPOT, 1982; Watanatada et al., 1987).

AR = the projected frontal area of the vehicle, in m^2, which may be user-specified or take on a default value as shown in Table 5.1.

Substituting these values in the force balance yields a cubic equation in VDRIVE. It may be shown that when the left hand term of the force balance, drive force, is positive the cubic equation always has a single positive root. Thus, given values of HPDRIVE and the other variables listed above, a unique VDRIVE value may be computed for a given vehicle type and road segment.

The absolute value of the vertical gradient, GR, depends on the average rise plus fall, RF, for the roadway and the sign of the gradient depends on whether it is on the uphill or downhill segment. Thus, solving the cubic equation with

$$GR = +\frac{RF}{1000}$$

would yield the value for $VDRIVE_u$, and solving it with

$$GR = -\frac{RF}{1000}$$

would yield the value for $VDRIVE_d$.

The value of the gross vehicle weight may be specified by the user as an option. Alternatively, it is computed endogenously as follows:

$$GVW = TARE + LOAD$$

where TARE = the vehicle tare weight, in tons, as given in Table 5A.1; and
LOAD = the vehicle payload, in tons, which may be user-specified or take on a default value as shown in Table 5.1.

The rolling resistance coefficient, CR, has been found empirically to be a function of road roughness, as given by:

$$CR = \begin{cases} 0.0218 + 0.0000467\ QI & \text{for cars and utilities} \\ 0.0139 + 0.0000198\ QI & \text{for buses and trucks} \end{cases}$$

where QI = the standard road roughness measure used in the Brazil-UNDP study, in QI units.

The mass density of air, in kilograms per cubic meter, RHO, is a function of altitude, given by (St. John et al., 1978):

$$RHO = 1.225\ [1 - 2.26 \times 10^{-5}\ A]^{4.255}$$

where A = the road altitude, defined as the elevation of the road section above the mean sea level, in meters.

VBRAKE

VBRAKE, the speed for a given road segment as limited by braking capacity and gradient, is arrived at using the concept of "used braking power," which is a positive quantity, denoted by HPBRAKE, in metric horsepower units. The assumption underlying the concept is that the steady-state speed attained on a long, smooth, straight downhill section is limited by the braking capacity, HPBRAKE, which depends on the vehicle type.

On an uphill segment the braking capacity constraint does not apply. Conceptually, when the brakes are not used the value of VBRAKE is infinity and 1/VBRAKE is zero. More generally, the constraint is not applicable whenever positive engine power is needed to impart motion to the vehicle. This would be the case on a downhill segment if the rolling resistance is greater in absolute value than the gradient-resistance; in symbols, whenver $CR \geq RF/1{,}000$.

When the constraint applies, VBRAKE and HPBRAKE are related, as before, through the force balance. However, since the braking capacity constraint is likely to become binding only for steep negative grades with low steady-state speeds, the air resistance term may be ignored without significant error. Thus the force balance may be solved as a first degree equation in VBRAKE.

The above remarks are summarized in these expressions:

$$VBRAKE_u \equiv \infty$$

$$VBRAKE_d = \begin{cases} \infty & \text{if } CR \geq RF/1000 \\ -\dfrac{736\ HPBRAKE}{g\ 1000\ GVW\ (CR - RF/1000)} & \text{if } CR < RF/1000. \end{cases}$$

HPBRAKE was estimated as a parameter using the speeds observed in the Brazil study. These values, appropriately modified, are given in Table 5.1 and may be used as defaults. Alternatively, the user has the option of substituting suitable values for a given practical application.

VCURVE

The curvature-limited speed, VCURVE, is arrived at from the postulate that when curvature is significant the speed is limited by the tendency of the wheels to skid. A good indicator of the tendency to skid is the ratio of the side force on the vehicle to the normal force. The used value of this ratio depends on the vehicle type and the surface type. It is called "perceived friction ratio", and is denoted by FRATIO. With some simplifications based on the facts that superelevation, SP, is small and that higher powers of small magnitudes are negligible, the ratio of forces is expressed by:

$$FRATIO = \frac{VCURVE^2}{g\ RC} - \frac{SP}{100}$$

whence $VCURVE = \sqrt{(FRATIO + SP/100)\ g\ RC}$.

The value of FRATIO has been derived as a function of the payload of the vehicle:

$$FRATIO = \max(0.02,\ FRATIO_0 - FRATIO_1\ LOAD)$$

where $FRATIO_0$ and $FRATIO_1$ are parameters which depend on the vehicle type as well as the surface type of the roadway. The values estimated from the Brazil data set are given in Table 5.1 and may be used as defaults. Optionally, the user may calibrate the parameters using the guidelines given in Ch. 13, Watanatada et al. (1987).

The radius of curvature, RC, is a simple function of average horizontal curvature:

$$RC = \frac{180{,}000}{\pi \max(18/\pi,\ C)}\ .$$

If the user does not supply values for superelevation, SP, the model computes values from the following formulas:

$$SP = \begin{cases} 0.012\ C & \text{for paved roads} \\ 0.017\ C & \text{for unpaved roads.} \end{cases}$$

These formulas are approximations to suggested design standards for typical speeds on these surfaces, and may be unrealistic for actual conditions in particular cases. Therefore, it is recommended that wherever possible the user provide values based on actual road geometry.

VROUGH

VROUGH is the limiting speed corresponding to the "maximum allowable" suspension motion of the vehicle, which governs ride severity. The ride suspension motion is measured by the rate of absolute displacements of the vehicle rear axle relative to the body, which, in turn, is related to vehicle speed and road roughness by the following relationship:

$$ARV = 0.0882\ V\ QI$$

where ARV = the average rectified velocity of suspension motion of the standard Opala-Maysmeter vehicle travelling at speed V, in mm/s;

V = the vehicle speed, in m/s; and

0.0882 = a constant for unit conversion.

The constraining speed due to ride severity, VROUGH, is assumed to be governed by the maximum allowable ARV, called ARVMAX, as follows:

$$VROUGH = \frac{ARVMAX}{0.0882\ QI}$$

where ARVMAX is an estimated parameter; estimates of ARVMAX for various vehicle classes are listed in Table 5.1.

VDESIR

Finally, VDESIR is the desired speed, i.e., the speed at which a vehicle is assumed to be operated in the absence of the constraints based on the vertical grade, curvature, and ride severity. The desired speed results from the driver's response to psychological, safety, economic, and other considerations. For each surface type (i.e., paved or unpaved), VDESIR was assumed to be constant for each vehicle type in the original Brazil study. The originally estimated values are listed in Table 5.1 as $VDESIR_0$ (in km/h). For narrow roads, however, VDESIR is assumed to be lower, and the following modified form, based on Indian data, has been adopted for use in HDM-III.

$$VDESIR = \begin{cases} VDESIR_0\ B_W/3.6 & \text{if effective number of lanes is 1.} \\ VDESIR_0/3.6 & \text{if effective number of lanes is more than 1 (that is, 1.5 or 2)} \end{cases}$$

where $VDESIR_0$ = the unmodified value of VDESIR obtained in the Brazil-UNDP study, in km/h;

B_W = dimensionless width parameter for adjusting desired speed.

In using the Brazil speed relationships the recommended ranges of the input variables shown in Table 5.2 should be observed to avoid extrapolating too far outside the range of the research. Local calibration of the model is always desirable, especially for the $VDESIR_0$, HPDRIVE and HPBRAKE parameters. Guidelines for adapting model parameters are given in Chapter 13 of Watanatada et al. (1987).

5.2.2 Fuel Consumption

Approach to fuel prediction

The fuel consumption prediction model makes use of the concept of the time-rate of fuel consumption or unit fuel consumption, denoted by UFC (in ml/s). Basic principles of interval combustion engine suggest that, under idealized environmental conditions, the unit fuel consumption is a function of power output (HP, in metric hp) and engine speed (RPM, in rpm). Symbolically, we may write:

UFC = UFC (HP, RPM).

While it is not possible to deduce the precise form of the UFC function from theoretical considerations, the function is known to be convex in both arguments. In the Brazil study, a quadratic form was employed, with separate coefficients for positive and negative power regimes. The estimated coefficients are given in Table 5.3a following the formal presentation of the model.

As in predicting vehicle speed, it has been found satisfactory to employ the idealized homogeneous uphill and downhill road segments for predicting fuel consumption. The power output may then be computed using the respective predicted speeds and the characteristics of the vehicles and the homogeneous segments.

In the Brazil experimental study, it was found that satisfactory prediction of fuel consumption could be obtained by using a constant nominal engine speed instead of actual used engine speed. The values of nominal engine speed (denoted by CRPM), as calibrated using the validation data from the Brazil study, are used as defaults in the fuel model. These are given in Table 5.3a. The user may supply individual values.

Since the test vehicles used in the Brazil study are makes and models typical in the mid-1970s, a "relative energy-efficiency factor," denoted by α_1, has been introduced to allow the model-user to incorporate changes in vehicle technology. This factor has a default value of 1.0 for makes and models close to the ones employed in the Brazil study. The user may specify lower values for newer, more fuel-efficient makes and models. Some typical values are presented in Table 5.3b.

Finally, in order to account for the differences between experimental conditions and real life driving conditions, "fuel adjustment factors," denoted by α_2, have been introduced. The values determined by calibrating the mechanistic fuel prediction models to the road user cost survey data are used as defaults in HDM-III. They are 1.16 for cars and utilities, and 1.15 for buses and trucks. The user may specify own values for α_2.

Full details regarding the estimation and validation of the fuel model may be found in Chapter 9 and 10 of Watanatada et al. (1987). Guidelines for local adaptation of CRPM, α_1. and α_2 parameters may be found in Chapter 14 of the same work.

Formal presentation

For a vehicle operating on any road section of specified geometric alignment, the average round trip fuel consumption, FL, in liters/1,000 vehicle-km is given by:

$$FL = 500\ \alpha_1\ \alpha_2\ \left[\frac{UFC_u}{V_u} + \frac{UFC_d}{V_d}\right]$$

where UFC_u = the predicted unit fuel consumption for the uphill segment, in ml/s;

UFC_d = the predicted unit fuel consumption for the downhill segment, in ml/s; and

α_1, α_2 are as explained earlier.

Table 5.2a: Recommended range of variables for speed, fuel consumption and tire wear prediction: Brazil relationships

Variable	Units	Recommended range
Road characteristics		
Altitude, A	m	0 - 5,000
Rise plus fall, RF	m/km	0 - 120
Horizontal curvature [1], C	degrees/km	0 - 1,200
Carriageway width, W	m	3.0 - 8.0
Superelevation, SP	%	0 - 20
Roughness, QI	QI	15 - 300
Vehicle fleet characteristics		
Gross vehicle weight, GVW	Tons	
Cars		0.8 - 2.0
Utilities		1.1 - 2.5
Large buses		7.5 - 12.0
Light trucks		3.0 - 6.5
Medium trucks		5.0 - 16.0
Heavy trucks		6.0 - 22.0
Articulated trucks		13.0 - 45.0
Payload, LOAD	Tons	
Cars		0 - 0.4
Utilities		0 - 1.4
Large buses		0 - 4.5
Light trucks		0 - 3.5
Medium trucks		0 - 11.0
Heavy trucks		0 - 16.0
Articulated trucks		0 - 32.0

[1] See text on tire wear prediction.
Source: Watanatada et al. (1987).

Table 5.2b: Recommended range of variables for speed, fuel consumption and tire wear prediction: Brazil relationships

Variable	Units	Recommended range
Vehicle fleet characteristics		
Maximum used driving power, HPDRIVE	Metric hp	
Cars		25 - 100
Utilities		35 - 100
Large buses		80 - 120
Light trucks		50 - 100
Medium trucks		80 - 120
Heavy trucks		80 - 120
Articulated trucks		180 - 230
Maximum used braking power, HPBRAKE	Metric hp	
Cars		15 - 30
Utilities		20 - 35
Large buses		140 - 180
Light trucks		90 - 120
Medium trucks		230 - 270
Heavy trucks		230 - 270
Articulated trucks		460 - 540
Projected frontal area, AR	m^2	
Cars		1.5 - 2.4
Utilities		2.3 - 3.2
Large buses		6.0 - 7.0
Light trucks		3.0 - 5.0
Medium trucks		5.0 - 8.0
Heavy trucks		5.0 - 8.0
Articulated trucks		5.5 - 10.0
Aerodynamic drag coefficient, CD	Dimensionless	0.3 - 1.0
Wearable rubber volume per tire, VOL	dm^3	
Large buses		5.6 - 8.0
Light trucks		2.0 - 3.5
Medium trucks		6.5 - 9.3
Heavy trucks		6.3 - 8.8
Articulated trucks		6.0 - 8.5

Source: Watanatada et al. (1987).

The predicted unit fuel consumption is given separately for the uphill and downhill road segments, as follows:

$$UFC_u = [UFC_0 + a_3\ HP_u + a_4\ HP_u\ CRPM + a_5\ HP_u^2\ ] \times 10^{-5}$$

$$UFC_d = \begin{cases} [UFC_0 + a_3\ HP_d + a_4\ HP_d\ CRPM + a_5\ HP_d^2] \times 10^{-5} & \text{if } HP_d \geq 0 \\ [UFC_0 + a_6\ HP_d + a_7\ HP_d^2] \times 10^{-5} & \text{if } NH_0 \leq HP_d < 0 \\ [UFC_0 + a_6\ NH_0 + a_7\ NH_0^2] \times 10^{-5} & \text{if } HP_d < NH_0 \end{cases}$$

where $UFC_0 = a_0 + a_1\ CRPM + a_2\ CRPM_2$;

CRPM = the calibrated engine speed, in revolutions per minute (default values of CRPM for the representative vehicles aregiven in Table 5.4, but the user may supply different values); and

HP_u, HP_d = the vehicle powers on the uphill and downhill road segments, in metric hp, given by:

$$HP_u = [(1000\ CR + RF)\ GVW\ g\ V_u + 0.5\ RHO\ CD\ AR\ V_u^3]/736$$

$$HP_d = [(1000\ CR - RF)\ GVW\ g\ V_d + 0.5\ RHO\ CD\ AR\ V_d^3]/736$$

and a_0 through a_7 and NH_0 are the parameters of the mechanistic fuel prediction models estimated using the data from controlled experiments. Their values are given in Table 5.3a.

In predicting fuel consumption using the Brazil relationships the recommended range of the input variables shown in Table 5.2 should be observed.

Figures 5.5 and 5.6 show predicted fuel consumption as a function of highway characteristics for a half-laden heavy truck over paved and unpaved roads, respectively. These curves have been drawn using the default values. Note that the fuel consumption curves incorporate the associated changes in speed induced by variation of the road characteristics.

5.2.3 Tire Wear

The HDM model employs two relationships obtained in the Brazil-UNDP study for predicting tire wear: one for cars and utilities, and another for trucks and buses. Because the tire data for cars and utilities obtained in the Brazil-UNDP study were inadequate, the relationship constructed with the data for these vehicle types is

Table 5.3a: Parameter values for fuel consumption prediction

Vehicle type	Small car	Medium car	Large car	Utility	Large bus	Light truck gas	Light truck diesel	Medium/ heavy truck	Articulated truck
CRPM (rpm)	3,500	3,000	3,300	3,300	2,300	3,300	2,600	1,800	1,700
a_τ	-8,201	23,453	-23,705	6,014	-7,276	-48,381	-41,803	-22,955	-30,559
a_1	33.4	40.6	100.8	37.6	63.5	127.1	71.6	95.0	156.1
a_2	0	0.01214	0	0	0	0	0	0	0
a_3	5,630	7,775	2,784	3,846	4,323	5,867	5,129	3,758	4,002
a_4	0	0	0.938	1.398	0	0	0	0	0
a_5	0	0	13.91	0	8.64	43.70	0	19.12	4.41
a_6	4,460	6,552	4,590	3,604	2,479	3,843	2,653	2,394	4,435
a_7	0	0	0	0	11.50	0	0	13.76	26.08
NH_τ	-10	-12	-15	-12	-50	-50	-30	-85	-85

Source: Watanatada et al. (1987).

relatively crude. On the other hand the more comprehensive data for trucks and buses permitted a more elaborate analysis based on mechanistic principles and idealized uphill and downhill road segments as in the speed and fuel relationships. Detailed description of these analyses are found in Chapter 11 of Watanatada et al., 1987. The relationships for computing the number of equivalent new tires consumed per 1,000 vehicle-kilometers, TC, for the various vehicle types are given as follows:

1. Passenger cars (small, medium and large) and utilities:

$$TC = \begin{cases} NT\ (0.0114 + 0.000137\ QI) & \text{for } 0 < QI \leq 200 \\ 0.0388\ NT & \text{for } QI > 200 \end{cases}$$

2. Light (gasoline and diesel), medium, heavy and articulated trucks and large buses:

$$TC = NT\left[\frac{(1 + 0.01\ RREC\ NR)\ [C_{otc} + C_{tcte}CF^2/L]}{(1 + NR)\ VOL} + 0.0075\right]$$

Table 5.3b: Relative energy-efficiency factors

Vehicle class	Test vehicle	Relative energy efficiency factor α_1		
		Comparable design	Modern design	Possible range
Small car	VW-1300	1.00	0.85	0.70-1.00
Medium car	Chevrolet Opala	1.00	0.85	0.70-1.00
Large car	Dodge Dart	1.00	0.95	0.80-1.00
Utility	VW-Kombi	1.00	0.95	0.80-1.00
Bus	Mercedes O-326	1.00	0.95	0.80-1.00
Light gasoline truck	Ford 400	1.00	0.95	0.80-1.00
Light diesel truck	Ford 4000	1.00	0.95	0.80-1.00
Medium truck	Mercedes 1113 (2 axles)	1.00	0.95	0.80-1.00
Heavy truck	Mercedes 1113 (3 axles)	1.00	0.95	0.80-1.00
Articulated truck	Scania 110/39	1.00	0.80	0.65-1.00

Source: Watanatada et al. (1987).

where NT = the number of tires per vehicle;

RREC = the ratio of the cost of one retreading to the cost of one new tire, in percent;

NR = the number of retreadings per tire carcass predicted by:

$NR = NR_o \exp(-0.00248\ QI - 0.00118\ C) - 1$;

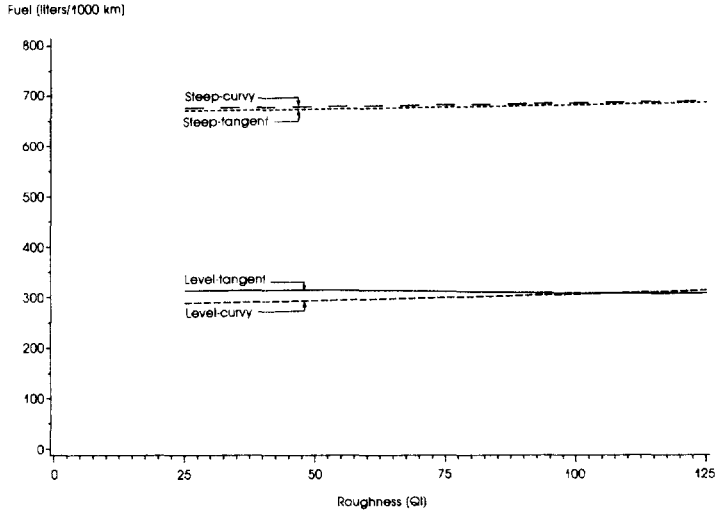
NR_o = the base number of recaps;

C_{otc} = the constant term of the tread wear model; and

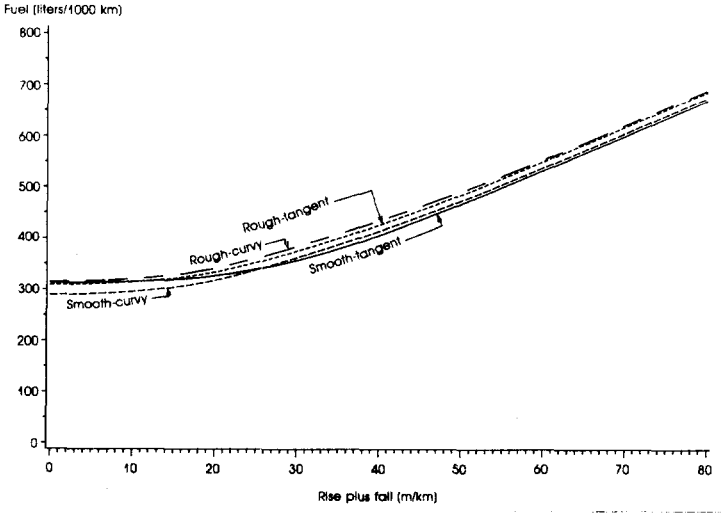
C_{tcte} = the wear coefficient;

Figure 5.5: Fuel as a function of road characteristics: half-laden heavy truck, paved road

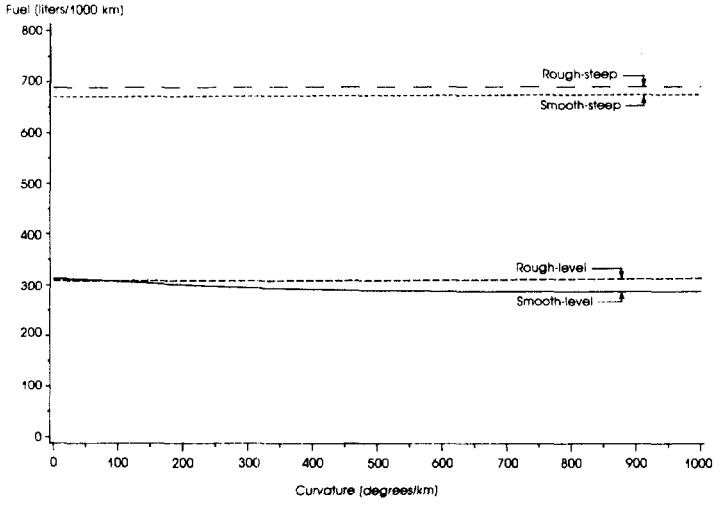
(a) Roughness



(b) Rise plus fall



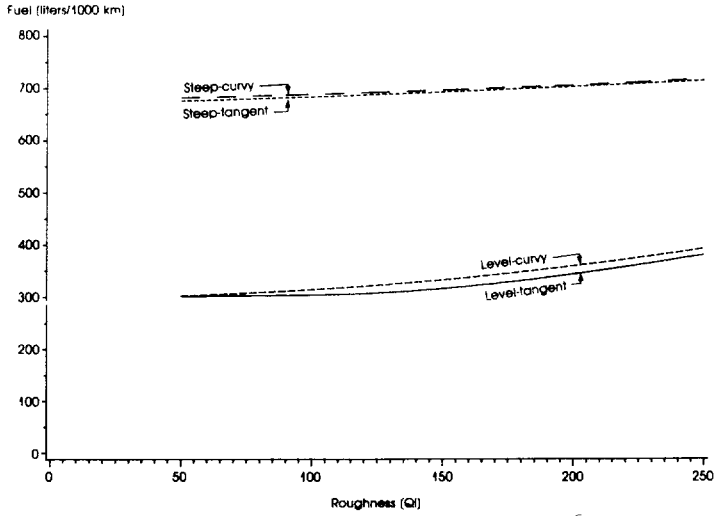
(c) Curvature



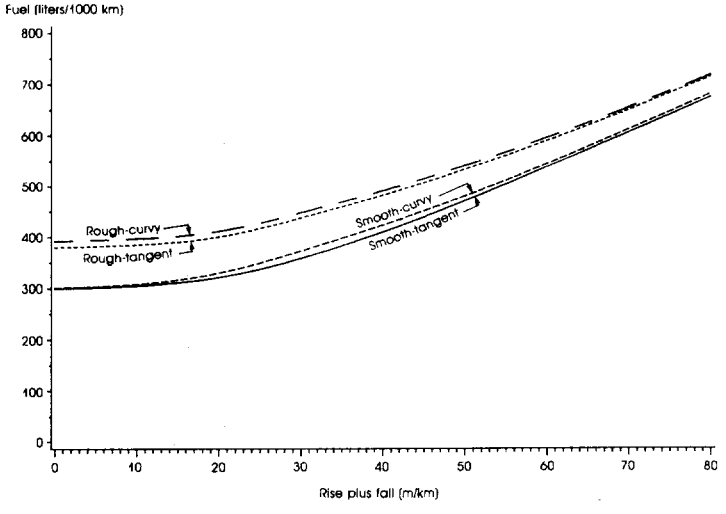
Source: Analysis of Brazil-UNDP-World Bank highway research project data (GEIPOT, 1982; Watanatada et al., 1987).

Figure 5.6: Fuel as a function of road characteristics: half-laden heavy truck, unpaved road

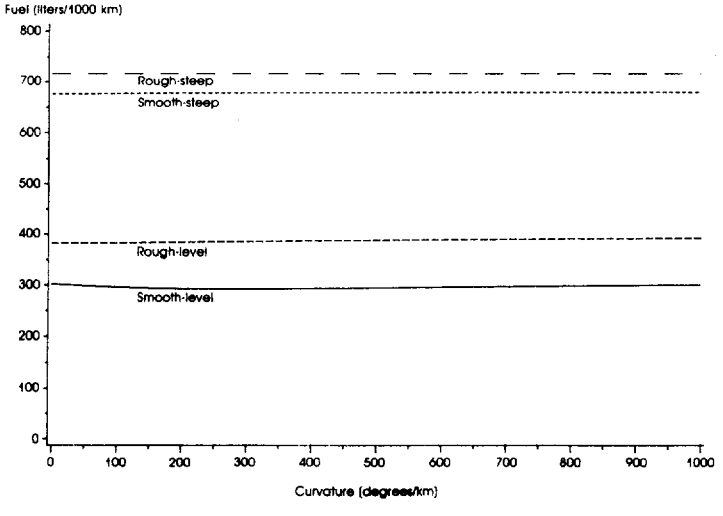
(a) Roughness



(b) Rise plus fall



(c) Curvature



Source: Analysis of Brazil-UNDP-World Bank highway research project data (GEIPOT, 1982; Watanatada et al., 1987).

CF^2 = the average squared circumferential force per tire, given by:

$$CF^2 = \frac{1}{2}(CF_u^2 + CF_d^2);$$

CF_u = the average circumferential force per tire (in the direction tangential to the road surface) on the uphill road segment, in newtons;

CF_d = the average circumferential force per tire (in the direction tangential to the road surface) on the downhill road segment, in newtons;

L = the average force per tire in the direction perpendicular to the road surface, in newtons, given by:

$$L = \frac{1000\ GVW\ g}{NT};$$

VOL = the average wearable rubber volume per tire for a given vehicle axle-wheel configuration and nominal tire size, in dm^3; and

0.0075 = the correction term for the prediction bias due to model nonlinearity.

Among the variables appearing in the formulas, NT, NR_o and VOL are parameters specific to vehicle types. Their default values are listed in Tables 5A.1 (for NT) and 5.4, but can be overridden by the user as an option. The parameters C_{otc} and C_{tcte} are specific, not to the vehicle class, but mainly to the material properties of the tire. The default values which apply to conventional (bias ply) type of tires of the Pirelli make are given in Table 5.4. RREC is a constant which has a default value of 15 percent representing the average for Brazil, but can also be overridden by the user.

The circumferential forces CF_u and CF_d are computed as the vehicle drive force (on the uphill and downhill segments, respectively) divided by the number of tires of the vehicle:

$$CF_u = \frac{1}{NT}[(1000\ CR + RF)\ GVW\ g + 0.5\ RHO\ CD\ AR\ V_u^2]$$

$$CF_d = \frac{1}{NT}[(1000\ CR - RF)\ GVW\ g + 0.5\ RHO\ CD\ AR\ V_d^2].$$

The above relationship is intended for use with roads of moderateorizontal alignment (C < 400 degrees/km) and well-designed superelevation. It is expected to underpredict the effect of horizontal alignment when the above conditions are not met. Furthermore the

Table 5.4: Parameter and default values for tire wear prediction

Vehicle types	NR_0	VOL (dm^3)	C_0 tc	C tcte (10^{-2})
Large bus	3.39	6.85	0.164	1.278
Light truck (diesel and gasoline)	1.93	4.30	0.164	1.278
Medium truck	3.39	7.60	0.164	1.278
Heavy truck	3.39	7.30	0.164	1.278
Articulated truck	4.57	8.39	0.164	1.278

Source: Watanatada et al. (1987).

recommended range of the input variables shown in Table 5.2 should be observed.

Figures 5.7 and 5.8 show predicted tire wear as a function of highway characteristics for a half-laden heavy truck over paved and unpaved roads, respectively. These curves have been plotted using the default values. It may be noted that the tire wear curves incorporate the associated changes in speed induced by variation of the road characteristics shown in Figures 5.3 and 5.4.

5.2.4 Maintenance Parts

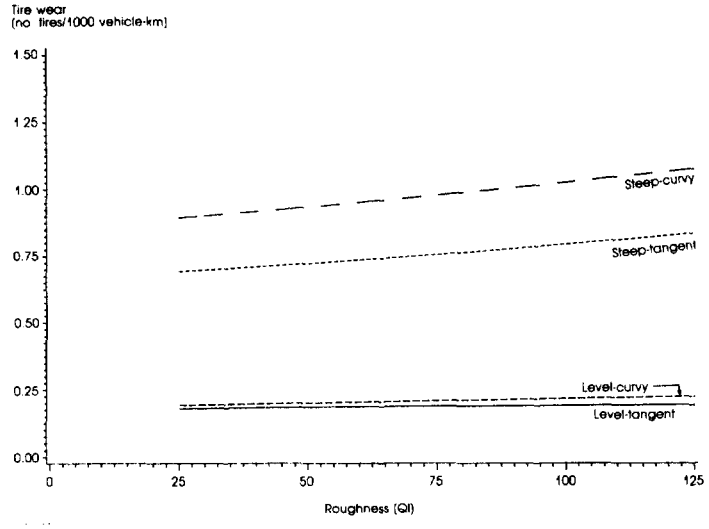
A convenient way of modeling parts consumption is in terms of the ratio of the monetary cost of parts consumed per 1000 vehicle-km to the price of a new vehicle in the same period. Under the assumption that the prices of spare parts and of a new vehicle vary together by the same proportions, data on their ratio from different periods may be included in the estimation and the results may be used to compute the monetary cost of spare parts consumed per 1000 vehicle-km for any year. For example, if the currency is Brazilian cruzeiros and the analysis year is 1990, then

$$MPC_{1990} = NVP_{1990}\ PC$$

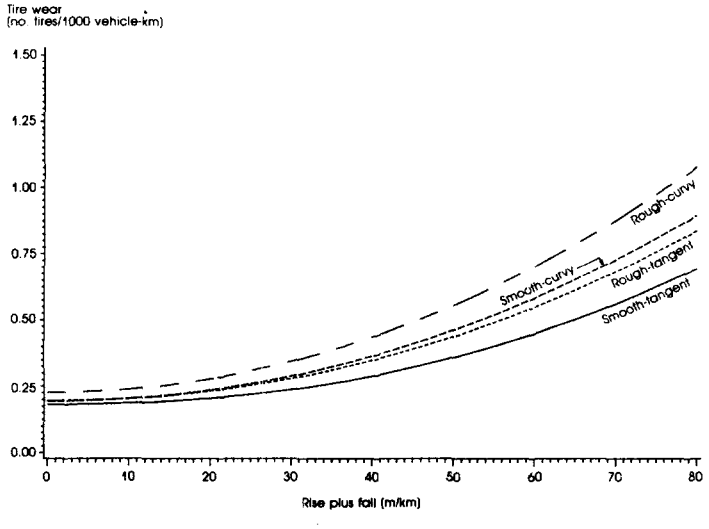
where MPC_{1990} = monetary cost of spare parts consumed, in 1990 Brazilian cruzeiros per 1,000 vehicle-km for the given vehicle class;

Figure 5.7: Tire wear as a function of road characteristics: half-laden heavy truck, paved road

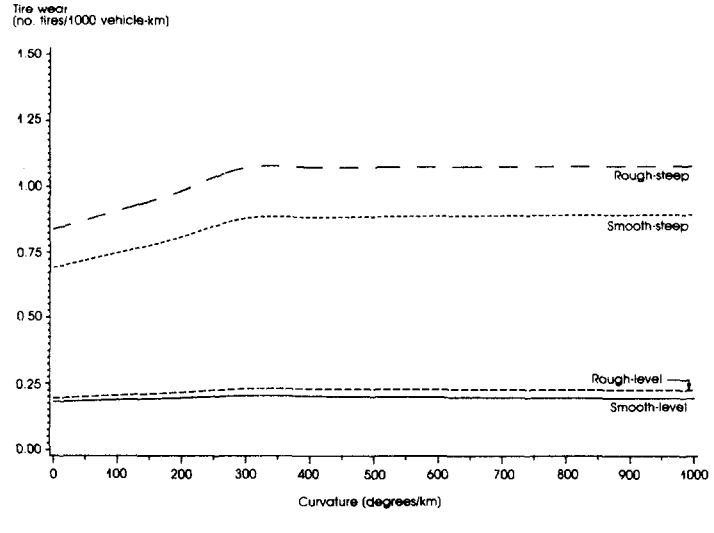
(a) Roughness



(b) Rise plus fall



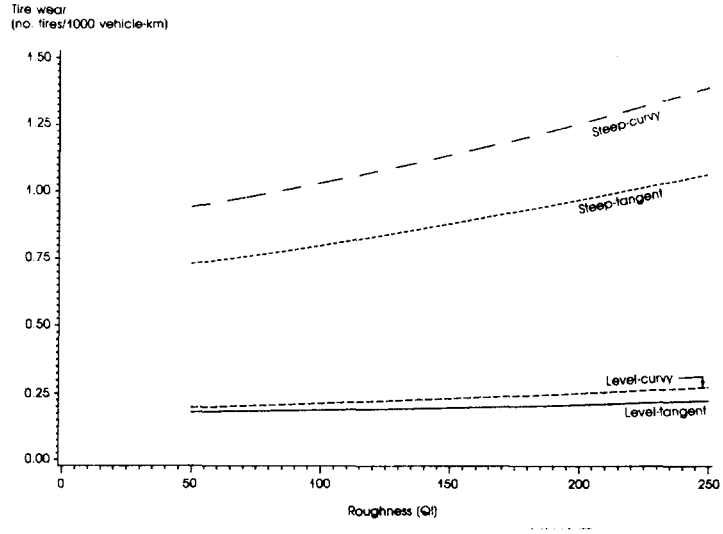
(c) Curvature



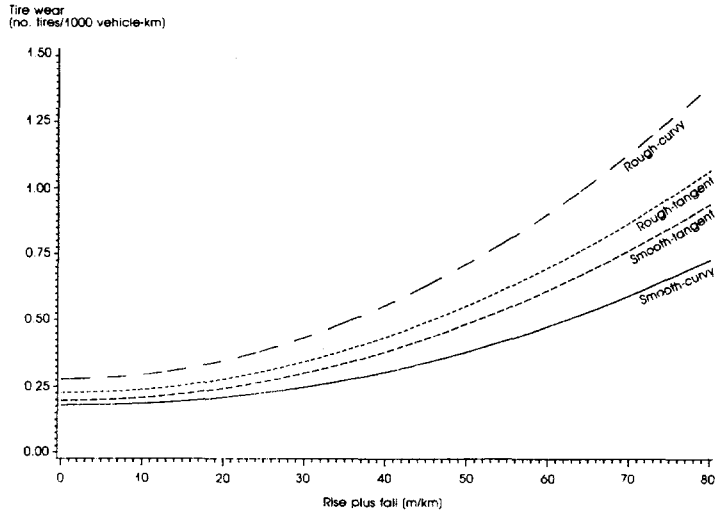
Source: Analysis of Brazil-UNDP-World Bank highway research project data (GEIPOT, 1982; Watanatada et al., 1987).

Figure 5.8: Tire wear as a function of road characteristics: half-laden heavy truck, unpaved road

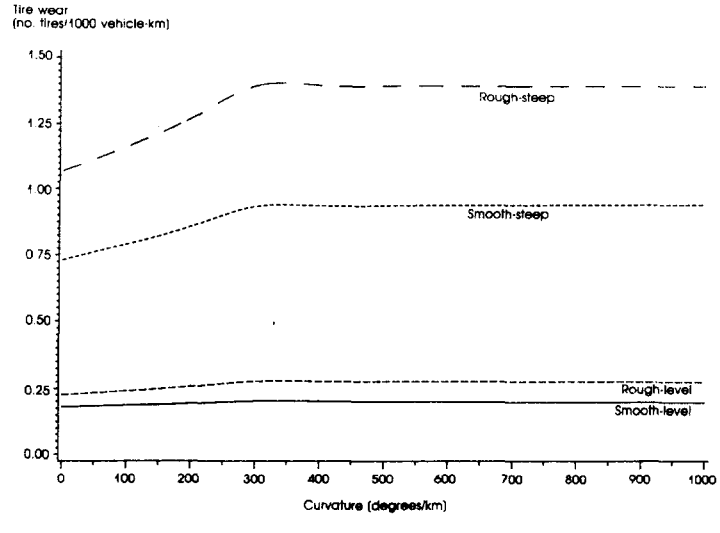
(a) Roughness



(b) Rise plus fall



(c) Curvature



Source: Analysis of Brazil-UNDP-World Bank highway research project data (GEIPOT, 1982; Watanatada *et al.*, 1987).

NVP_{1990} = expected average price of a new vehicle of the given class, in 1990 Brazilian cruzeiros; and

PC = the parts cost per 1,000 vehicle-km for the given vehicle class expressed as a fraction of the cost of a new vehicle.

The rest of this subsection is devoted to discussion of the equations to predict PC.

PC has been found to be related to roughness and vehicle age (Chesher and Harrison, 1987). The effects of these two factors are multiplicative. Holding the age constant, the relationship between PC and roughness is generally exponential, especially for relatively low values of roughness. However, the exponential relation tends to overpredict PC at higher values. Therefore, the recommended equation is a composite of exponential and linear -- exponential up to a transitional value of roughness, QI_{osp}, which is different for different vehicle types, and then linear for higher values. The linear extension is tangent to the exponential relationship at QI_{osp}.[3] Since the Brazil relation for truck parts consumption was found to be linear over all values of roughness encountered in practice, QI_{osp} is set to zero for all trucks.

The average age of the vehicles in a region belonging to a given vehicle group is expressed by the average cumulative kilometrage of the group (denoted by CKM). The effect of CKM has been found to be multiplicative with an exponent. The age exponent (denoted by "k") is remarkably stable for a given vehicle type over environments as varied as Brazil and India.

Symbolically, PC is given by:

$$PC = \begin{cases} CKM^k \, C_{osp} \exp(C_{spqi} \, QI) & \text{for } QI < QI_{osp} \\ CKM^k \, (a_o + a_1 QI) & \text{for } QI > QI_{osp} \end{cases}$$

where CKM = the average age of the vehicle group in km, defined as the average number of kilometers the vehicles have been driven since they were built;

k = the age exponent -- a fixed model parameter (given in Table 5.5);

C_{osp} = the constant coefficient in the exponential relationship between spare parts consumption and roughness -- an optional user-input model parameter (with default values);

---

[3] The use of a tangential extension for extrapolation of the log-linear model was introduced by A. Chesher (GEIPOT, 1982).

C_{spqi} = the roughness coefficient in the exponential relationship between spare parts consumption and roughness -- an optional user-input model parameter (with default values);

$QI_{o\,sp}$ = the transitional value of roughness, in QI, beyond which the relationship between spare parts consumption and roughness is linear -- an optional user-input model parameter (with default values).

a_0 and a_1 are coefficients of the linear-tangential extension of the exponential relationship and may be expressed as functions of model parameters. The relationships are:

$$a_o = C_{o\,sp}\exp(C_{spqi}\ QI_{o\,sp})(1 - C_{spqi}\ QI_{o\,sp})$$

$$a_1 = C_{o\,sp}\ C_{spqi}\ \exp(C_{spqi}\ QI_{o\,sp})$$

A convenient formula to arrive at a good estimate of CKM is

$$CKM = \min\ (1/2\ LIFE_o AKM_o,\ CKM^1)$$

where $LIFE_o$ = the average vehicle service life in years, input by the user, and

Table 5.5: Parameters and default values for predicting spare parts spare parts consumption

Vehicle classes	k	$C_{o\,sp}$ (10^{-6})	C_{spqi} (10^{-3}/QI)	$QI_{o\,sp}$ (QI)	CKM' (km)
Car and utility	0.308	32.49	13.70	120	300,000
Large bus	0.483	1.77	3.56	190	1,000,000
Light (gasoline and diesel) and medium truck	0.371	1.49	251.79	0	600,000
Heavy truck	0.371	8.61	35.31	0	600,000
Articulated truck (tractor and trailer)	0.371	13.94	15.65	0	600,000

Source: Chesher and Harrison (1987), modified for unit changes and the tangential extension. The parameters for utilities given there are not used because of the differences in the associated makes and models.

AKM_o = the average number of kilometers driven per year per vehicle of the group, input by the user, and

CKM' = the ceiling on average cumulative kilometerage given in Table 5.5.

The subscript "$_o$" is used to emphasize that the values of these variables are to be provided by the user and are not the same as the values calculated in terms of speeds for depreciation and interest calculations (Sec. 5.3.3). The ceiling on average cumulative kilometrage is introduced because high values of CKM could otherwise lead to unrealistically high predictions of parts consumption.

The prediction of PC requires three parameters, namely C_{osp}, C_{spqi} and QI_{osp}. These should be provided by the user from local data, but default values will be used if inputs are not provided. The default values, based on the Brazil study are given in Table 5.5. The recommended ranges for the independent variable QI, for use with the default parameters, are given in Table 5.6.

Figure 5.9 shows maintenance parts consumption as a function of roughness for different vehicle classes. These curves have been drawn using the default values.

5.2.5 Maintenance Labor

Maintenance labor hours are related primarily to maintenance parts requirements, and in some cases, to roughness. When significant, the latter has been found to be exponential and the two effects are multiplicative. The relationship in its general form is written as:

$$LH = C_{o\ell h}\, PC^{C_{\ell hpc}} \exp(C_{\ell hqi} QI)$$

where LH = the predicted number of maintenance labor-hours per 1,000 vehicle-km;

PC = the standardized parts cost per 1,000 vehicle-km expressed as a fraction of new vehicle price;

$C_{o\ell h}$ = the constant coefficient in the relationship between labor hours and parts costs, which is an optional user-input model parameter;

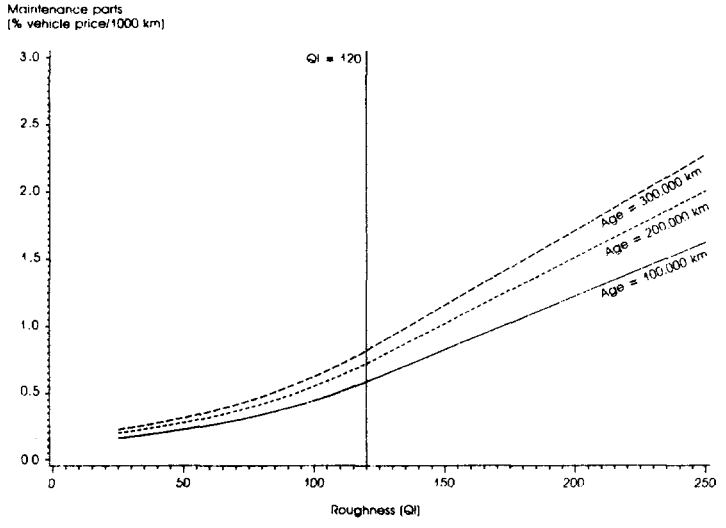
$C_{\ell hpc}$ = the exponent of parts cost in the relationship between labor hours and parts cost, which is an optional user-input model parameter; and

$C_{\ell hqi}$ = the roughness coefficient in the exponential relationship between labor hours and roughness, which is an optional user-input model parameter.

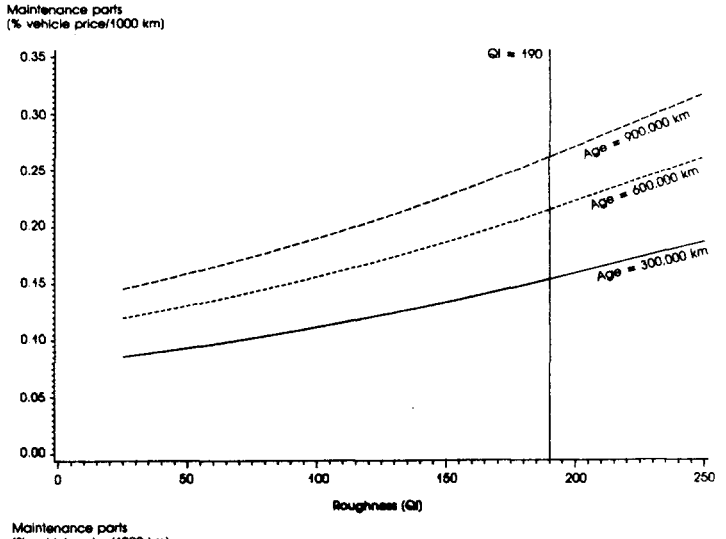
The prediction of LH requires three parameters, namely, $C_{o\ell h}$, $C_{\ell hpc}$, and $C_{\ell hqi}$. These should be provided by the user, but default

Figure 5.9: Maintenance Parts Consumption as a Function of Road Roughness: Various Vehicles

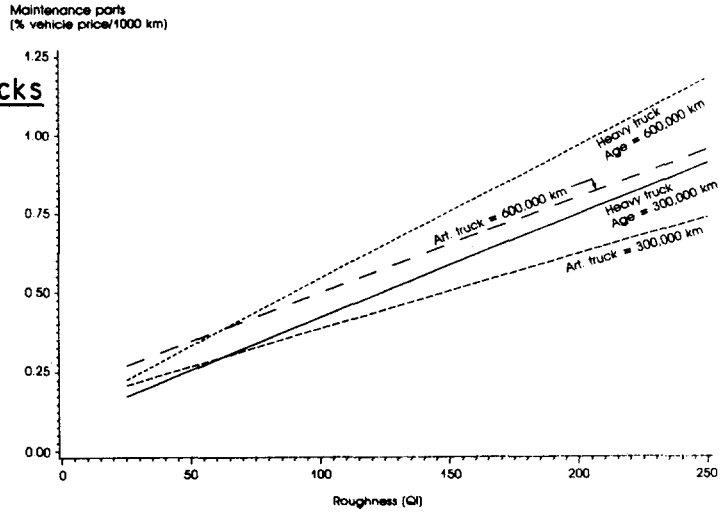
(a) Cars and utilities



(b) Large buses



(c) Heavy and articulated trucks



Source: Analysis of Brazil-UNDP-World Bank highway research project (GEIPOT, 1982; Chesher and Harrison, 1987, Watanatada et al., 1987).

Table 5.6: Recommended range of variables for maintenance parts and labor prediction

Variable	Units	Recommended range
Roughness, QI	QI	
Cars and utilities		25 - 120
Large buses		25 - 190
Trucks		25 - 120

Source: Based on data from GEIPOT (1982, Volume 5).

values will be used if inputs are not provided. The default values, estimated from the Brazil data are given in Table 5.7. The recommended range for the independent variable QI is the same as for predicting parts requirements (Table 5.6).

Figure 5.10 shows maintenance labor hours needed as a function of roughness for different vehicle classes. These curves have been drawn using the default values.

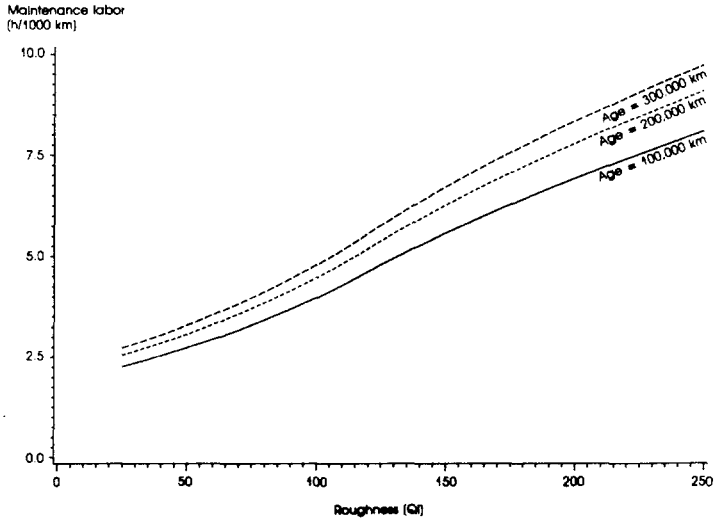
Table 5.7: Default values of parameters for computing maintenance labor-hour prediction

Vehicle classes	$C_{\ell h}$	$C_{\ell hpc}$	$C_{\ell hqi}$ (QI^{-1})
Car and utility	77.14	0.547	0
Large bus	293.44	0.517	0.0055
Light (gasoline and diesel) and medium truck	242.03	0.519	0
Heavy truck	301.46	0.519	0
Articulated truck (tractor and trailer)	652.51	0.519	0

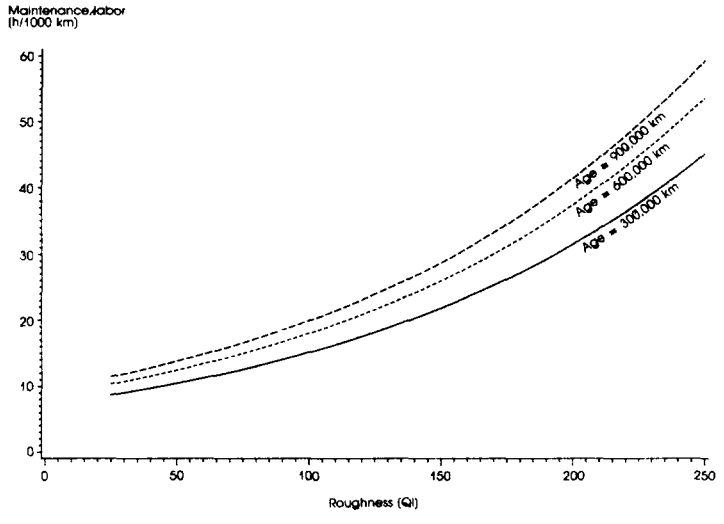
Source: Chesher and Harrison (1985), modified for unit changes. The parameters for utilities given there are not used because of the differences in makes and models.

Figure 5.10: Maintenance labor hours as a function of road roughness: various vehicles

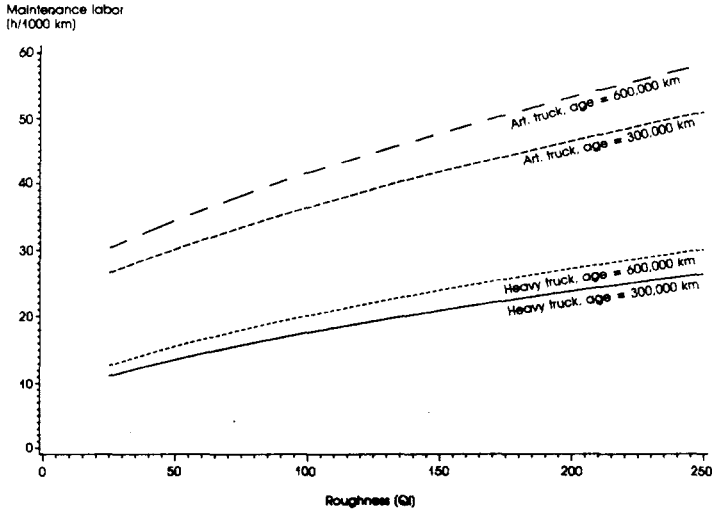
(a) Cars and utilities



(b) Large buses



(c) Heavy and articulated trucks



Source: Analysis of Brazil-UNDP-World Bank highway research project (GEIPOT, 1982; Chesher and Harrison, 1987, Watanatada et al., 1987).

5.3 RELATIONSHIPS USED WITH ALL OPTIONS

The portion of the vehicle operating cost submodel explained thus far is the portion for which different formulations are available from the four different studies. It includes relationships from the Brazil-based option for vehicle speed, fuel consumption, tire wear, maintenance parts, and, maintenance labor. The alternative sets of relations for these elements from the other country studies are presented and explained in the next chapter.

Relations for other components involved in vehicle operating cost are not differentiated according to the different sources, but are the same regardless of which option is chosen for the elements listed above. These other components are lubricants consumption, crew, depreciation, interest, overhead, passenger delays, cargo holding, and miscellaneous costs.

Relations for these components are described in the rest of this chapter.

5.3.1 Lubricants Consumption

Lubricants consumption is predicted as a function of roughness using relationships modified by Chesher and Harrison (1985), based on the India study (CRRI, 1982):

1. Passenger cars and utilities
   $OC = 1.55 + 0.000211\ BI$

2. Light trucks
   $OC = 2.20 + 0.000211\ BI$

3. Buses and medium and heavy trucks
   $OC = 3.07 + 0.000211\ BI$

4. Articulated trucks
   $OC = 5.15 + 0.000211\ BI$

where OC = the lubricants consumption, in liters/1,000 vehicle-km.

5.3.2 Crew Requirements

In the HDM model the cost of crew labor is considered to be a variable cost rather than a fixed cost. This means that the time the crew spends on non-driving activities such as loading, unloading, and layovers is not charged against this cost category. Thus, the number of crew hours required per 1,000 vehicle-km, denoted by CRH, is inversely proportional to speed:

$$CRH = \frac{1000}{S}.$$

5.3.3 Vehicle Depreciation and Interest

The amount of vehicle depreciation per 1,000 vehicle-km, denoted by DEP and expressed as a fraction of the average cost of a new vehicle, is computed as:

$$DEP = 1000 \frac{ADEP}{AKM}$$

where ADEP = the average annual vehicle depreciation, expressed as a fraction of the average cost of a new vehicle; and

AKM = the average number of kilometers driven per vehicle per year.

Similarly, the amount of interest charge per 1,000 vehicle-km, denoted by INT and expressed as a fraction of the average new vehicle cost, is given by:

$$INT = 1000 \frac{AINT}{AKM}$$

where AINT = the average annual interest on the vehicle expressed as a fraction of the average cost of a new vehicle.

To calculate depreciation and interest costs per 1,000 vehicle-km, two steps are required: first, determine the average annual depreciation and interest, and, second, determine the average annual kilometerage. Two optional methods are available in the HDM model for computing the average annual depreciation (ADEP) and interest (AINT):

1. de Weille's varying vehicle life method.

2. Constant vehicle life method.

The above methods correspond to the DEPRECIATION codes in the input data (User's Manual, Chapter 2, Form D-5).

Similarly, three optional methods are available in the model for computing the average annual utilization (km):

1. Constant annual kilometerage method,
2. Constant annual hourly utilization method, and
3. Adjusted utilization method;

where the first two methods are, in fact, special cases of the more general third method. These three methods correspond to the UTILIZATION codes in the User's Manual, Chapter 2 (Form D-5). The DEPRECIATION and UTILIZATION methods can be specified for individual vehicle groups in any combination. The following paragraphs provide detailed descriptions of these methods.

Average annual depreciation and interest

1. de Weille's varying vehicle life method. This method, suggested by de Weille (1966), is based on straight-line depreciation over a predetermined vehicle service life which is assumed to decrease somewhat as vehicle speed increases, so that lifetime kilometerage increases in less proportion than speed. Mathematically, de Weille's method is expressed as:

Annual depreciation factor: $ADEP = \frac{1}{LIFE}$

Annual interest factor: $AINT = \frac{AINV}{2} \bullet \frac{1}{100}$

where AINV = the annual interest charge on the purchase cost of a new vehicle, in percent; and

LIFE = the average vehicle service life, in years.

The vehicle service life is assumed to be related to the predicted vehicle operating speed, S, as follows:

$$LIFE = \min \left\{1.5\ LIFE_0;\ \left[\frac{S_0}{S} + 2\right] \frac{LIFE_0}{3}\right\}$$

where $LIFE_0$ = the baseline average vehicle life, in years, specified by the user; and

S_0 = the baseline average vehicle speed, in km/h, given by:

$$S_0 = \frac{AKM_0}{HRD_0}$$

where AKM_0 = the user-specified baseline average annual kilometerage (as defined before); and

HRD_0 = the user-specified baseline number of hours driven per vehicle per year.

The maximum limit of 1.5 $LIFE_0$ is imposed in the model to eliminate the possibility of computing unrealistically long vehicle lives. In addition to the method for estimating vehicle life, de Weille (1966) also suggested a method for estimating the average annual kilometerage driven per year, called the "constant hourly utilization method" to be described subsequently.

2. Constant vehicle life method. This method also uses straight-line depreciation, but differs from de Weille's method in that the vehicle life, LIFE, is assumed to be constant irrespective of vehicle speed and equal to the user-specified value, i.e.,

$$LIFE = LIFE_0.$$

The interest factor, being independent of vehicle life, is the same as in de Weille's method.

Average annual utilization

1. Constant annual kilometerage method. This method assumes that for each vehicle group the average annual kilometerage driven per vehicle, AKM, is constant and equal to the user-specified value, i.e.:

$$AKM = AKM_0.$$

The assumption of constant annual kilometerage may be considered appropriate for non-commercial vehicles such as private passenger cars, of which the trip distances and frequencies are usually assumed to be relatively insensitive to changes in average travel speed. However, commercial vehicles tend to be used for more frequent or longer trips if time for a given length trip is reduced.

2. Constant annual hourly utilization method. This method, which was also suggested by de Weille (1966), assumes that the average annual number of hours driven per vehicle is constant. Thus the average annual kilometerage driven per vehicle, AKM, is computed as the product of the user-specified average number of hours driven per vehicle per year, HRD_0, and speed, S:

$$AKM = HRD_0 \, S.$$

It should be noted that the constant hourly utilization method tends to overpredict time-related benefits. As the average speed increases the number of trips a commercial vehicle can make in a year tends to rise. However, in addition to the driving time the total time needed to complete each trip also includes a large proportion of non-driving activities such as loading, unloading, vehicle services and repairs, and layovers. This means that the number of trips per year does not increase in direct proportion to speed; doubling the speed, for example, results in less than doubling the number of trips and annual kilometerage. The effect is particularly pronounced for short trips with many stops for pickups and deliveries.

3. Adjusted utilization method. This method aims to remedy the deficiencies in both of the other methods. Each vehicle is assumed in this method to operate on the section under consideration as a fixed route over the analysis year, with the total time, per round trip, TT, given by:

$$TT = TN + TD$$

where TN = the time spent on non-driving activities as part of the round trip tour, including loading and unloading, refueling, layovers, etc., in hours per trip; and

TD = the driving time on the section, in hours per trip.

Letting RL denote the round trip driving distance or route length, in km, the driving time on the section, TD, becomes:

$$TD = \frac{RL}{S}$$

It can be seen that as the operating speed, S, increases, the driving time per trip decreases. Vehicle operators are assumed to try to make as many trips as possible within the total number of hours available per year, denoted by HAV. The latter is an operating constraint which depends on the time allowed for crew rest, infeasibility of vehicle operation (e.g., during early morning, vehicle repairs, etc.). HAV is assumed to be independent of the vehicle speed and the route characteristics. Under this assumption the kilometerage driven per year is derived as:

$$AKM = \text{(number of round trips per year) x (route length)}$$

$$= \frac{HAV}{TT} RL = \frac{HAV \quad RL}{TN + \frac{RL}{S}}$$

$$= \frac{HAV}{\frac{TN}{RL} + \frac{1}{S}} .$$

The term TN/RL, the number of non-driving hours per vehicle-km of travel, can be expressed as follows:

$$\frac{TN}{RL} = \frac{HAV - HRD_o}{AKM_o}$$

Substituting this expression into the above expression for AKM yields:

$$AKM = \frac{HAV}{\frac{HAV - HRD_o}{AKM_o} + \frac{1}{S}} .$$

Thus if the values of HAV, HRD_0 and AKM_0 are available, the annual kilometerage driven, AKM, can be predicted as a function of the predicted operating speed, S.

The baseline parameters HRD_0 and AKM_0 are to be provided directly by the user. The HAV parameter is derived from the following formula:

$$HAV = \frac{HRD_o}{HURATIO_o}$$

where $HURATIO_o$ = the "hourly utilization ratio" for the baseline case. Substituting this formula in the expression for AKM above yields the general formula for predicting vehicle utilization:

$$AKM = \frac{AKM_o\ HRD_o}{HRD_o\ (1 - HURATIO\ ) + \frac{AKM_o\ HURATIO_o}{S}} .$$

Based on the data from the Brazil-UNDP study, Watanatada et al. (1987) estimated typical values of hourly utilization ratio for various classes of vehicles. These results led to the following default values of $HURATIO_o$ in the HDM model.

Vehicle type	Baseline hourly utilization ratio ($HURATIO_o$)
Cars	0.60
Utilities	0.80
Buses	0.75
Trucks	0.85

These default values are based on the operating characteristics of typical vehicles in Brazil. However, since hourly utilization ratios are expected to vary considerably across countries, it is important to provide values appropriate to the local operating conditions.

The behavior of the average annual kilometerage driven, AKM, with respect to the hourly utilization ratio, $HURATIO_0$, is of interest. On one extreme, when $HURATIO_0$ equals zero, the above general formula reduces to:

$$AKM = AKM_0$$

which is recognized as the "constant annual kilometerage" method. On the opposite extreme, when $HURATIO_0$ equals one, the formula becomes:

$$AKM = HRD_0 \, S$$

which is recognized as the "constant annual hourly utilization" method.

Figures 5.11 and 5.12 show depreciation and interest costs as a function of highway characteristics for a half-laden heavy truck for paved and unpaved roads, respectively. These curves have been drawn using the default values. The method used to predict vehicle utilization is the adjusted utilization method.

5.3.4 Overhead

Two optional methods are available for computing the overhead cost of vehicle operation for each vehicle group:

1. As a lump sum overhead cost per vehicle prorated over the annual kilometerage, AKM; and

2. As a percentage of the vehicle running costs (consisting of the costs of fuel and lubricants consumption, tire wear, maintenance parts and labor, depreciation and interest and crew).

Only one method may be used for each vehicle group.

5.3.5 Passenger Delays

The number of passenger-hours spent in traveling per 1000 vehicle-km, denoted by PXH, is given by:

$$PXH = 1000 \frac{PAX}{S}$$

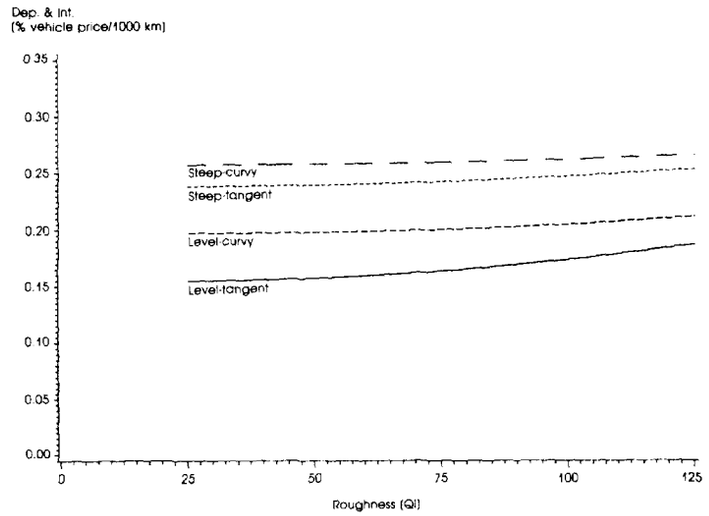
where PAX = the average number of passengers per vehicle, input by the user.

5.3.6 Cargo Holding

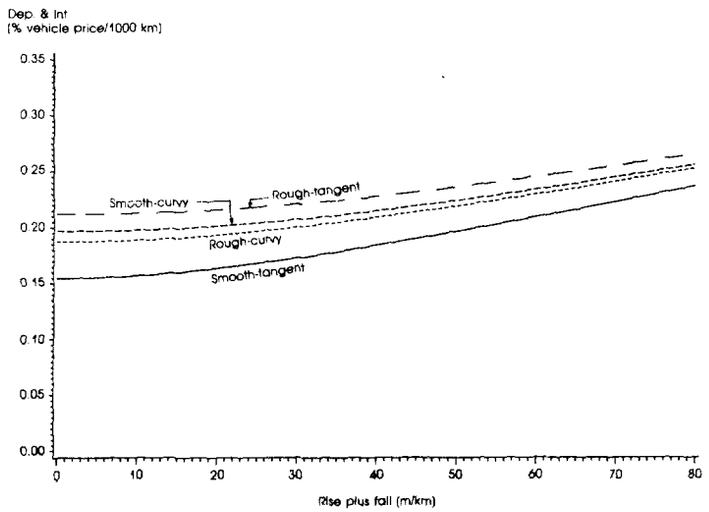
The number of vehicle-hours spent in transit per 1000 vehicle-km, denoted by VCH, is given by:

Figure 5.11: Depreciation and interest as a function of road characteristics: half-laden heavy truck, paved road

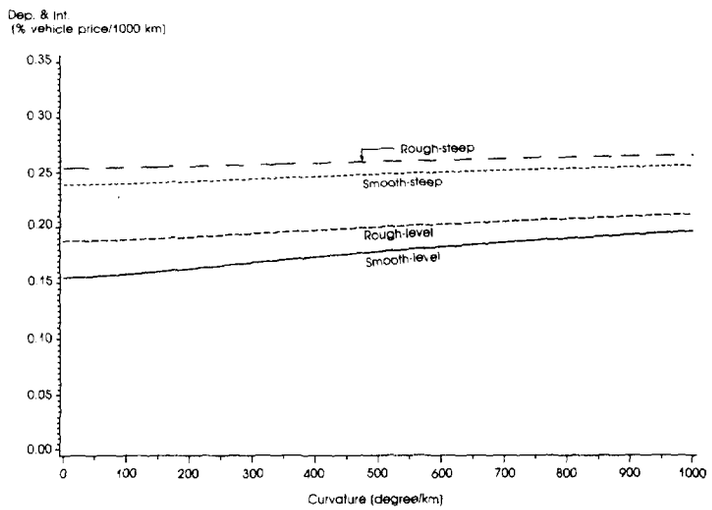
(a) Roughness



(b) Rise plus fall



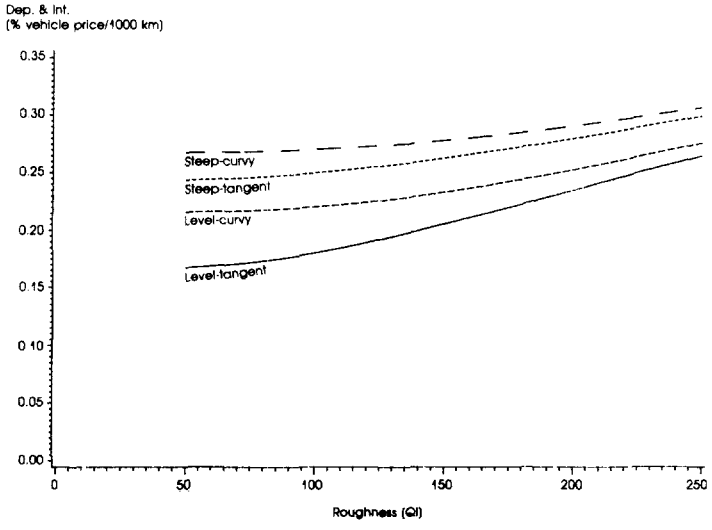
(c) Curvature



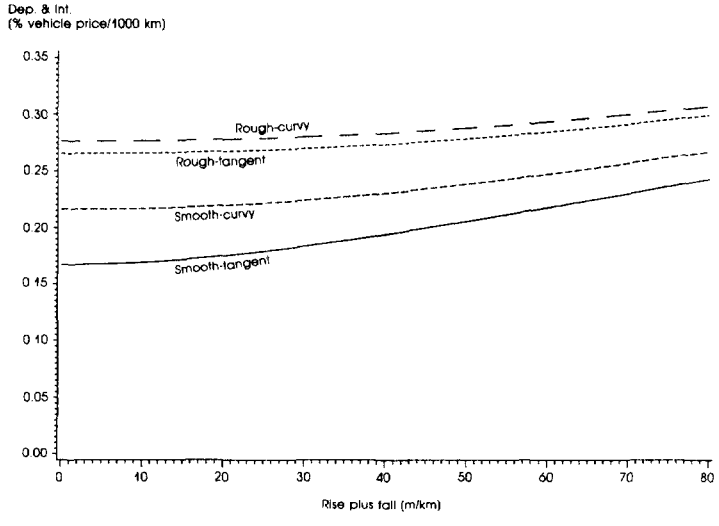
Source: Analysis of Brazil-UNDP-World Bank highway research project (GEIPOT, 1982; Chesher and Harrison, 1987, Watanatada et al., 1987).

Figure 5.12: Depreciation and interest as a function of road characteristics: half-laden truck, unpaved road

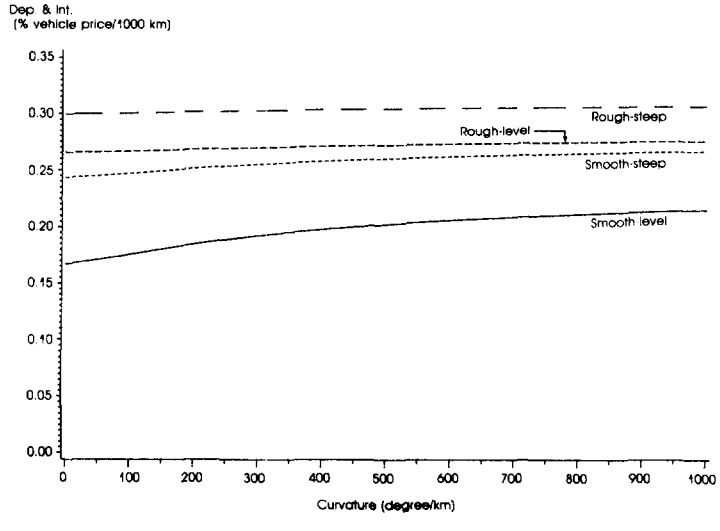
(a) Roughness



(b) Rise plus fall



(c) Curvature



Source: Analysis of Brazil-UNDP-World Bank highway research project (GEIPOT, 1982; Chesher and Harrison, 1987, Watanatada et al., 1987).

$$VCH = \frac{1000}{S} .$$

The cargo holding cost per 1000 vehicle-km is defined as the product of VCH and the user-specified cargo holding cost per vehicle-hour delayed. If time savings or delays are perfectly deterministic and there are no problems with lumpiness in scheduling, the unit cargo holding cost per vehicle-hour may be computed as equal to:

$$\frac{\text{(the monetary value of cargo) x (the annual interest rate, in fraction)}}{\text{8760 hours per year}}$$

If time savings or delays are stochastic and scheduling lumpiness prevails (as is normally the case at least to some extent), the unit cargo holding cost will generally be higher, due to the need to build up extra inventories to avoid stocking out. In general, cargo holding cost is much smaller than other vehicle operating cost components and therefore, although it is included in the HDM model, it is often ignored for simplicity in HDM applications.

5.3.7 Miscellaneous Costs

In the preceeding discussion relating various vehicles operating cost components to road geometry and surface conditions, it was assumed that average roughness of the roadway was a sufficient statistic for all the road condition factors affecting the cost of operating vehicles on the roadway. It is obviously an over-simplification, because road condition factors -- such as, moisture, depth of loose material, and reduced passability due to gravel loss -- are relevant for vehicle operating costs, but are not captured by roughness. Of such factors, impassability of an unpaved road due to insufficient gravel "cover" is the most important. Hence, provision has been made in the model to incorporate the expected rise in the vehicle operating costs once the gravel thickness drops below a minimum level.

Increase in vehicle operating costs can be attributed to:

1. Increased tire slippage in loose (e.g., sandy) or soft (e.g., muddy) materials;

2. Increased drag in losse or soft material;

3. Increased torque requirements and consequent wear on the vehicle drive assembly; and

4. Reduced travel speed, including probably forced stops where the vehicle may partially or completely bog down; etc. None of these are reflected directly in the measured roughness data.

Rather than quantify these effects physically, their economic impact on the road user's selection of vehicle may be considered. For example, if the road is rendered impassable for significant periods, the road user would select a four-wheel drive vehicle or truck at greater costs in order to gain passability. These greater costs give a reasonable estimate of the cost of impassability and thus the benefits of passability.

The cost of marginal passability are modelled by a piecewise linear function, FPASS, of gravel thickness, which augments the unit vehicle operating costs up to a limit, denoted by FPLIMIT. This limit is specified explicitly by the user based on an understanding of the nature and cause of passability-loss (e.g., by saturation, loss of cohesion [in sands], etc.) The function FPASS would begin with a value of 1 when the gravel thickness (GH) equalled the minimum thickness required (GHMIN) and increase linearly to the limiting value (FPLIMIT) when the thickness reached nil.

The treatment of impassability costs, <u>for a given vehicle class</u>, may be formalized as follows:

$$CPASS = VOC\ (FPASS - 1)$$

where CPASS = cost due to impassability, in monetary units per 1,000 km;

VOC = total unit vehicle operating costs computed as a sum of components discussed so far, in monetary units per 1000 km;

FPASS = dimensionless factor derived from gravel thickness, given by

$$FPASS = 1 + (FPLIMIT - 1)\ \max\ (0;\ 0 - GH/GHMIN);$$

FPLIMIT = vehicle-specific dimensionless maximum value of FPASS, specified by the user;

GH = mean gravel thickness for the year, in mm; and

GHMIN = minimum gravel thickness, in mm, determined as

$$GHMIN = \min\ (100.0;\ \max\ (40.0;\ 2.0\ D95));\ \text{where}$$

D95 = maximum particle size, in mm.

Note that the physical explanation for increasing FPASS for gravel thickness less than the minimum is that there is greater risk of weak spots, and of increased vehicle costs, in this range than for roads with adequate gravel "cover" thickness.

The factor FPLIMIT ranges in value from 1 for subgrade materials with soaked CBR greater than 10 percent (i.e., fully passable) to 3 for heavy vehicles on soft soils. A default value of 1.0 is used which the user may override. By definition, CPASS is zero for a paved road, and GH is zero for an earth road.

5.4 UNIT COSTS

Up to this point the analysis has dealt, wherever feasible, with physical quantities of resources used, so that fundamental physical relations would not be obscured by price variations. Once determined, the physical quantities are multiplied by unit costs or prices in a separate step. The unit costs or prices must be provided by the user.

Physical concepts proved difficult to define and quantify for maintenance parts and for overhead and miscellaneous costs, and are not relevant for depreciation and interest, which are financial in nature. For three of these elements, i.e., all except overhead and miscellaneous costs, it was found convenient and valid to deal with the ratio of the element's cost to the price of a new vehicle, another cost factor to be supplied by the user. Overhead is treated, at the user's option, either as an exogenous lump sum or as a percentage of running costs. Miscellaneous costs are also treated as a fraction of other costs.

Table 5.8 shows the units in which each element of resource consumption is measured and the dimensions of the price, unit cost, or other factor by which each has to be multiplied to obtain its value as a component of vehicle operating cost per 1,000 vehicle-kilometers. Multiplying these values by the number of thousand vehicle-kilometers traveled on the link in question by each vehicle group in a year and adding the group totals yields the total vehicle operating cost on the link for the year.

Table 5.8: Costing of vehicle operating resources

Component	Units of measurement	Unit cost or other multiplying factor
Fuel consumption	liters/1,000 vehicle-km	Cost per liter
Tire wear	Equiv. new tires per 1,000 vehicle-km	Cost per tire
Maintenance parts	Proportion of new vehicle cost per 1,000 vehicle-km	Cost of new vehicle
Maintenance labor	Labor-hours/1,000 veh.-km	Wage cost per hour
Lubricant use	Liters per 1,000 veh.-km	Cost per liter
Crew time	Person-hrs./1,000 veh.-km	Crew cost per hour
Depreciation	Fraction of new vehicle cost per 1,000 veh.-km	Cost of new vehicle
Interest	Fraction of new vehicle cost per 1,000 veh.-km	Cost of new vehicle
Overhead	1. lump sum + annual km	No factor needed:
	or 2. % of running costs	Sum of costs above
Passenger time	Pass.-hours/1,000 veh.-km	Value per hr. of passenger time
Cargo holding	Vehicle-hrs./1,000 veh.-km	Cargo holding cost per vehicle-hour
Miscellaneous costs	Fraction of above costs	Sum of costs above

APPENDIX 5A

COMPARATIVE VEHICLE CHARACTERISTICS AND INPUT REQUIREMENTS FOR ALTERNATIVE VEHICLE OPERATING COST MODELS

Table 5A.1: Vehicle characteristics

Vehicle type[1]	Approx. tare weight (tons)	Approx. rated gross weight (tons)	Weight classification[2]	Tires	Heavy axles[3]	Engine: Fuel[4] type	Engine: Maximum SAE rated power (metric hp)	Engine: Cylinders	Representative test fleet vehicle
Brazil[5]									
1. Passenger cars (small)	1.0	1.2	L	4	0	G	49	4	VW 1300
2. Passenger cars (medium)	1.2	1.5	L	4	0	G	148	6	Opala
3. Passenger cars (large)	1.7	1.9	L	4	0	G	201	8	Dodge Dart
4. Utilities	1.3	2.1	L	4	0	G	61	4	VW Kombi
5. Buses	8.1	11.5	H	6	2	D	149	6	Merc. Benz 0362
6. Light trucks (gasoline)	3.1	6.1	H	6	2	G	171	8	Ford-400
7. Light trucks (diesel)	3.3	6.1	H	6	2	D	103	4	Ford-4000
8. Medium trucks	5.4	15.0	H	6	2	D	149	6	Merc. Benz 1113[6]
9. Heavy trucks	6.6	18.5	H	10	3	D	149	6	Merc. Benz 1113[7]
10. Articulated trucks	14.7	40.0	H	18	5	D	289	6	Scania 110/39
India									
1. Passenger cars (small)	0.9	1.2	L	4	0	G	49	4	Premier Padmini car
2. Passenger cars (medium)	1.2	1.5	L	4	0	G	51	4	Ambassador car
3. Utilities	0.6	1.2	L	4	0	G	39	4	Mahindra jeep
4. Medium trucks	6.1	12.2	H	6	2	D	114	6	Tata truck
5. Heavy trucks	8.1	16.3	H	6	2	D	127	6	Ashok truck
6. Buses (Tata)	-	11.0	H	6	2	D	114	6	None
7. Buses (Ashok)	-	13.0	H	6	2	D	127	6	None
Kenya									
1. Passenger cars	-	0.9	L	4	0	G	72	4	Ford Cortina estate car
2. Utilities	-	1.7	L	4	0	G	87	6	Land-Rover
3. Light trucks	3.3	8.3	H	6	2	D	109	6	Bedford truck
4. Buses	-	10.0	H	-	2	D	140	-	None
5. Medium trucks	-	13.0	H	-	2	D	140	-	None
Caribbean									
1. Passenger cars	-	1.1	L	4	0	G	65	4	Ford Cortina estate car
2. Utilities	1.6	2.6	L	4	0	G	71	4	Ford Transit van
3. Light trucks	4.0	8.4	H	6	2	D	115	6	Ford truck D1010

1 Figures are those of representative vehicles.
2 L = light vehicle rated gross weight under 3.5 [metric] tons; H = heavy vehicle [rated gross weight of 3.5 tons or more].
3 The number of heavy axles equals the number of axles of the vehicle if it is a heavy vehicle and zero otherwise.
4 G = gasoline engine; D = diesel engine.
5 Tare weights of Brazil vehicles include 150 kg weight of two drivers.
6 Excludes third rear axle.
7 Includes third rear axle.

Source: Kenya: Hide et al. (1975); Caribbean: Hide (1982); Morosiuk and Abayanayaka (1982); Brazil: GEIPOT (1982); Watanatada et al.(1987); India: Central Road Research Institute (1982).

Table 5A.2: Vehicle fleet characteristics data for vehicle operating cost submodel

Variable	Study region			
	Brazil	Kenya	Caribbean	India
Maximum rated power (metric hp)[1]		0	0	0
Gross vehicle weight (tons)[2]	0	0	0	0
Number of tires per vehicle[3]	0			0
Vehicle payload (tons)	0			
Used driving power (metric hp)	0			
Used braking power (metric hp)	0			
Aerodynamic drag coefficient (dimensionless)	0			
Projected frontal area (m^2)	0			
Calibrated engine speed (rpm)	0			
Wearable rubber volume per tire (dm^3)	0			
Vehicle service life (years)[4]	R	R	R	R
Annual number of km driven[4]	R	R	R	R
Annual number of hours driven[4]	R	R	R	R
Hourly utilization ratio (dimensionless)[5]	0	0	0	0

Note: R = required input; 0 = optional (a default value provided in the model for each vehicle type can be overridden by user-specified value).

[1] Relevant only for trucks and buses.

[2] Not relevant for cars and utilities in the Kenya, Caribbean and India relationships.

[3] Not relevant for cars and utilities in the India relationships.

[4] Generally required with few exceptions.

[5] Defined as the ratio of the number of hours driven to the number of hours the vehicle is available.

Table 5A.3: Road and environmental characteristics data for vehicle operating cost submodel

Variable	Study region			
	Brazil	Kenya	Caribbean	India
Environment				
Altitude (m)	0	R		
Road geometry and surface type and conditions				
Rise plus fall (m/km)	R	R	R	R
Horizontal curvature (degrees/km)	R	R	R	R
Carriageway width (m)	R	R	R	R
Superelevation (percent)	0			
Surface type (paved/unpaved)	R	R		
Roughness	R	R	R	R

Note: R = required input; 0 = default values are provided in the model which can be overridden by the user as an option.

CHAPTER 6

Alternative Vehicle Operating Cost Relationships

6.1 SELECTION OF THE RELATIONS

The first two-thirds of Chapter 5 presented the principal set of relations for predicting vehicle speed, fuel consumption, tire wear, and maintenance parts and labor, which are based on data collected in the GEIPOT-UNDP-IBRD study in Brazil. Except for the maintenance components, these were formulated according to mechanistic and behavioral principles. The present chapter presents several alternative sets of relations which were developed from data in Kenya, Caribbean and India, respectively, and were analyzed by other methods.

As discussed in Chapter 5, the user is advised normally to use the Brazil relationships with as much local calibration as possible. The principal exceptions would be for applications in India, Kenya, and the Caribbean countries where the alternative relationships were themselves statistically established, and even there the user is encouraged to use the Brazil relationships also in parallel analyses and to compare the results in order to gain a more thorough understanding of the likely effects of alternative road policies. For a comparative evaluation of the alternative sets of relationships, including their graphical representations, see the first volume in the HDM series (Chesher and Harrison, 1987).

All of the different sets of relations are included in the HDM-III model, and the user must specify which set is to be used. The same set will be used in its entirety throughout any run. Input data requirements differ somewhat for different sets, and the data required by the set chosen must, of course, be supplied. Appendix 5A presents comparative vehicle characteristics and input requirements for all sets of relationships. Regardless of which set of relations is chosen for the cost components mentioned above, the relations for other components of vehicle operating cost -- lubricants, crew time, depreciation, interest, overhead, passenger time, cargo holding and miscellaneous costs -- remain the same as those presented in the last part of Chapter 5.

The alternative sets of relationships are essentially empirical in nature, and thus have limited extrapolative validity beyond the range of the independent variables in the estimation data. These ranges are given for each components, based either on the recommendations in the respective study, as in the case of the Kenya and Caribbean relationships, or on the actual range in the data used for estimation, as in the case of the India relationships. As such, the ranges for a given independent variable are not always consistent across different cost components in the same study. It is recommended that the user pay close attention to the ranges of independent variables and discrepancies within them by greater scrutiny of

the detailed user cost and resource consumption reports available as an output option in applications of the HDM model.

The rest of this chapter describes the relationships derived from the Kenya, Caribbean, and India studies for estimating vehicle speeds, fuel consumption, tire wear, and maintenance parts and labor.

6.2 THE KENYA RELATIONSHIPS

The relationships described in this section were derived from the TRRL study in Kenya (Hide et al., 1975) with the modifications described in Appendix 6A.

6.2.1 Vehicle Speed

The average operating speeds of vehicles plying the road are required for estimating fuel consumption, the unit costs per 1000 vehicle-km of crew, depreciation and interest, passenger delays, and cargo holding. Speeds are calculated as a function of the average rise plus fall, horizontal curvature, carriageway width, altitude, roughness, and, for trucks and buses, power-to-weight ratio.

The round-trip journey speed, S, in km/h, is defined as:

$$S = \frac{2}{\frac{1}{S_u} + \frac{1}{S_d}}$$

where S_u = speed on the uphill segment of the trip, in km/h and
S_d = speed on the downhill segment of the trip, in km/h.

The expressions for S_u and S_d are as follows:

Unpaved roads

1. **Passenger cars**

$$S_u = \max [20; 86.7 - 0.209\ RF - 0.118\ C - 0.00089\ BI - 0.00491\ A - 4.32 \max (0; 5 - W)]$$
$$S_d = \max [20; 86.7 - 0.070\ RF - 0.118\ C - 0.00089\ BI - 0.00491\ A - 4.32 \max (0; 5 - W)]$$

2. **Utilities**

$$S_u = \max [20; 80.3 - 0.317\ RF - 0.0966\ C - 0.00095\ BI - 0.00278\ A - 4.32 \max (0; 5 - W)]$$
$$S_d = \max [20; 80.3 - 0.059\ RF - 0.0966\ C - 0.00095\ BI - 0.00278\ A - 4.32 \max (0; 5 - W)]$$

3. **Buses**

$$S_u = \max [15; 65.9 - 0.492\ RF - 0.0463\ C - 0.00036\ BI - 0.00417\ A - 6.36 \max (0; 5 - W)]$$
$$S_d = \max [156\ 65.9 - 0.010\ RF - 0.0463\ C - 0.00036\ BI - 0.00417\ A - 6.36 \max (0; 5 - W)]$$

4. **Light and medium trucks**

$$S_u = \max [15;\ 44.1 - 0.433\ RF - 0.061\ C - 0.00060\ BI - 0.00042\ A + 1.10\ PWR - 6.36 \max (0;\ 5 - W)]$$

$$S_d = \max [15;\ 44.1 + 0.00445\ RF - 0.061\ C - 0.00060\ BI - 0.00042\ A + 1.10\ PWR - 6.36 \max (0;\ 5 - W)]$$

Paved roads

1. **Passenger cars**

$$S_u = \max [20;\ 105.3 - 0.372\ RF - 0.110\ C - 0.00089\ BI - 0.00491\ A - 7.31 \max (0;\ 5 - W)]$$

$$S_d = \max [20;\ 105.3 - 0.0759\ RF - 0.110\ C - 0.00089\ BI - 0.00491\ A - 7.31 \max (0;\ 5 - W)]$$

2. **Utilities**

$$S_u = \max [20;\ 89.7 - 0.418\ RF - 0.0738\ C - 0.00095\ BI - 0.00278\ A - 7.31 \max (0;\ 5 - W)]$$

$$S_d = \max [20;\ 89.7 - 0.0496\ RF - 0.0738\ C - 0.00095\ BI - 0.00278\ A - 7.31 \max (0;\ 5 - W)]$$

3. **Buses**

$$S_u = \max [15;\ 73.4 - 0.526\ RF - 0.0661\ C - 0.00036\ BI - 0.00417\ A - 3.29 \max (0;\ 5 - W)]$$

$$S_d = \max [15;\ 73.4 + 0.0666\ RF - 0.0661\ C - 0.00036\ BI - 0.00417\ A - 3.29 \max (0;\ 5 - W)]$$

4. **Light and medium trucks**

$$S_u = \max [15;\ 49.8 - 0.519\ RF - 0.0581\ C - 0.00060\ BI - 0.00042\ A + 1.10\ PWR - 3.29 \max (0;\ 5 - W)]$$

$$S_d = \max [15;\ 49.8 + 0.030\ RF - 0.0581\ C - 0.00060\ BI - 0.00042\ A + 1.10\ PWR - 3.29 \max (0;5 - W)]$$

where
- RF = the average rise plus fall, in m/km (see definition in 5.2.1);
- C = the average horizontal curvature, in degrees/km (see definition in 5.2.1);
- BI = the road roughness, in mm/km, as measured by the TRRL Towed Fifth Wheel Bump Integrator (Hide et al., 1975);
- W = the average width of the carriageway, in meters;
- A = the road altitude, defined as the elevation of the road section above the mean sea level, in meters; and
- PWR = the vehicle power-to-weight ratio, given by:
- PWR = HPRATED/GVW

where
- HPRATED = the maximum rated power of the engine, in metric hp, which may be user-specified or take on the default value shown in Table 5A.1; and
- GVW = the gross vehicle weight, in (metric) tons, which may be user-specified or take on the default value shown in Table 5A.1 as rated gross vehicle weight.

Note that one metric hp is defined as equal to 75 kg-m/s; this makes one metric hp equal to 0.986 British hp.

In the TRRL's analysis of speed on paved roads, the effect of roughness had been found statistically insignificant. This is judged to be the result of an insufficiently wide range of roughness in the test roads. Assuming that a given change in roughness would have the same effect with or without paving, a term based on unpaved road results has been added to the equations for speed on paved roads to account for roughness. The incorporation of this term and an offsetting constant are explained in Appendix 6A, Section 6A.1.

Because of the linear form of the original speed equations the value of speed computed can possibly be unrealistically low or even negative. To circumvent this possibility, minimum limits, which roughly correspond to crawl speeds on very poor surfaces or alignments, have been imposed, i.e., 20 km/h for cars and utilities and 15 km/h for trucks and buses, for both paved and unpaved roads. The user should observe the recommended range of the input variables used in predicting speeds recommended by TRRL, as compiled in Table 6.1, as deviation from this range represents an extrapolation and could produce unreasonable results.

6.2.2 Fuel Consumption

Two optional sets of relationships are provided for estimating fuel consumption for paved and unpaved roads, one set obtained from the original TRRL-Kenya study (Hide et al., 1975) and the other from subsequent work by Chesher (1977a) based on the same set of data.

The original TRRL-Kenya relationships express fuel consumption, through multiple linear regression equations, in terms of the vehicle

Table 6.1: Recommended range of variables for vehicle speeds predictions: Kenya relationships

Variable	Units	Recommended range
Unpaved roads		
Rise plus fall, RF	m/km	0-120
Horizontal curvature, C	Degrees/km	0-250
Carriageway width, W	m	3.0-8.0
Roughness, BI	mm/km	2,000-14,000
Altitude	m	0-2,500
Paved roads		
Rise plus fall, RF	m/km	0- 120
Horizontal curvature, C	Degrees/km	0- 200
Carriageway width, W	m	3.0- 8.0
Roughness, BI	mm/km	2,000-9,000
Altitude, A	m	0-2,500
Vehicle fleet characteristics		
Power-to-weight ratio [1], PWR	Metric hp/ton	13 - 33

[1] For light and medium trucks.
Source: Based on data from Hide et al. (1975).

speed, average rise and fall, gross vehicle weight, power-to-weight ratio and, for unpaved roads, roughness and looseness of surface material. These relationships were slightly modified, eliminating looseness as an explicit variable (by incorporating the mean value of 1.0 mm), and converting from British to metric horsepower.

The predicted round-trip fuel consumption, FL, in liters/1000 km is defined as

$$FL = \frac{FL_u + FL_d}{2}$$

where FL_u = fuel consumption on the uphill segment of the trip, in liters/1,000 km; and

FL_d = fuel consumption on the downhill segment of the trip, in liters/1,000 km

The expression for FL_u and FL_d are as follows:

Unpaved roads

1. **Passenger cars**

$$FL_u = 1.16\,[47.7 + 614/S_u + 0.0079\,S_u^2 + 1.723\,RF + 0.0011\,BI]$$

$$FL_d = 1.16\,[47.7 + 614/S_d + 0.0079\,S_d^2 - 1.066\,RF + 0.0011\,BI]$$

2. **Utilities**

$$FL_u = 1.16\,[74.6 + 844/S_u + 0.0137\,S_u^2 + 2.828\,RF + 0.0011\,BI]$$

$$FL_d = 1.16\,[74.6 + 844/S_d + 0.0137\,S_u^2 - 1.306\,RF + 0.0011\,BI]$$

3. **Light trucks**

$$FL_u = 1.15\,[124.0 + 796/S_u + 0.0150\,S_u^2 + 4.176\,RF + 0.0014\,BI - 2.58\,\min(PWR;\,40)]$$

$$FL_d = 1.15\,[124.0 + 796/S_d + 0.0150\,S_d^2 + 2.216\,RF + 0.0014\,BI - 2.58\,\min(PWR;\,40)]$$

4. **Medium trucks and buses**

$$FL_u = 1.15\,[-30.04 + 796/S_u + 0.0150\,S_u^2 - 4.176\,RF + 0.0015\,BI + \max(-2.58\,PWR + 69.2\,\sqrt{GVW}\,;\,0)]$$

$$FL_d = 1.15\ [-30.04 + 796/S_d + 0.0150\ S_d^2 - 2.216\ RF + 0.0015\ BI + \max(-2.58\ PWR + 69.2\ \sqrt{GVW}\ ;\ 0)]$$

Paved roads

1. **Passenger cars**

$$FL_u = 1.16\ [53.4 + 499/S_u + 0.0058\ S_u^2 + 1.594\ RF + DFL_R]$$

$$FL_u = 1.16\ [53.4 + 499/S_d + 0.0058\ S_d^2 - 0.854\ RF + DFL_R]$$

2. **Utilities**

$$FL_u = 1.16\ [74.7 + 1151/S_u + 0.0131\ S_u^2 + 2.906\ RF + DFL_R]$$

$$FL_d = 1.16\ [74.7 + 1151/S_d + 0.0131\ S_d^2 - 1.277\ RF + DFL_R]$$

3. **Light trucks**

$$FL_u = 1.15\ [105.4 + 903/S_u + 0.0143\ S_u^2 + 4.362\ RF - 2.37\ \min(PWR;\ 40) + DFL_R]$$

$$FL_d = 1.15\ [105.4 + 903/S_d + 0.0143\ S_d^2 - 1.834\ RF - 2.37\ \min(PWR;\ 40) + DFL_R]$$

4. **Medium trucks and buses**

$$FL_u = 1.15\ [-48.6 + 903/S_u + 0.0143\ S_u^2 + 4.362\ RF + \max(-\ 2.37\ PWR + 69.2\ \sqrt{GVW};\ 0) + DFL_R]$$

$$FL_d = 1.15\ [-48.6 + 903/S_d + 0.0143\ S_d^2 - 1.834\ RF + \max(-\ 2.37\ PRW + 69.2\ \sqrt{GVW};\ 0) + DFL_R]$$

where FL = the predicted fuel consumption, in liters/1,000 vehicle-km; and

DFL_R = the predicted increase in fuel consumption due to road roughness, in liters/1,000 vehicle-km, to be described subsequently; and the other variables are as defined before.

Multiple regression relationships were originally obtained from controlled field experiments. To extend the resulting equations to real-world operating conditions, the multiplicative factors, derived from the Brazil study, 1.16 for cars and utilities and 1.15 for trucks and buses, are used.

As in the speed relationships, the effect of roughness on fuel consumption on paved roads was found by TRRL to be statistically insignificant. Again it was assumed that this finding resulted from the narrowness of the range of roughness, and an adjustment based on unpaved road results was added to the fuel consumption equations. As with the speed relationships, a constant was included to make the calculated result match that of the TRRL in the middle of the observed range. The expressions for the adjustment, DFL_R, for different vehicle types are:

$$DFL_R = \begin{cases} -3.3 + 0.0011\ BI & \text{for passenger cars and utilities} \\ -4.2 + 0.0014\ BI & \text{for trucks and buses} \end{cases}$$

The constants 3.3 and 4.2 have been chosen so that at the roughness level of 3,000 mm/km, which is approximately the mean roughness value for the Kenya test paved road sections, the adjustment DFL_R becomes zero. In the equations for buses and trucks limits have been imposed on the power-to-weight ratio (PWR) and gross vehicle weight (GVW) terms to ensure that fuel consumption predictions do not become negative.

In using the fuel consumption relationships the user should observe the recommended range of the input variables shown in Table 6.2, as deviation from the range represents an extrapolation.

6.2.3 Alternative Kenya Fuel Consumption Relationships

Based on the same set of field data as the original TRRL-Kenya equations, Chesher (1977a) developed a new set of fuel consumption relationships which incorporate non-linear effects of the independent variables, notably the rise plus fall and the vehicle power-to-weight ratio and operating speed, as well as the interactions of these variables. These relationships, which are offered as an option in the model, are described in detail in Appendix 6A.2

6.2.4 Tire Wear

Based on the road user survey data with road roughness ranging from 2,500 to 7,500 mm/km, TRRL (Hide et al., 1975) obtained the following relationships for tire consumption:

1. **Passenger cars and utilities**

$$TC = \begin{cases} [-83 + 0.058\ BI]\ 10^{-3} & BI \geq 2{,}000\ mm/km \\ 0.03 & BI < 2{,}000\ mm/km \end{cases}$$

2. **Light and medium trucks and buses**

$$TC = \begin{cases} GVW\ [83 + 0.0112\ BI]\ 10^{-4} & BI \geq 1{,}500\ mm/km \\ 0.01\ GVW & BI < 1{,}500\ mm/km \end{cases}$$

where TC = the number of cost-equivalent new tires consumed per 1,000 vehicle-km.

Table 6.2: Recommended range of variables for fuel consumption prediction: Kenya relationships

Variable	Units	Recommended range
Unpaved roads		
Rise plus fall, RF	m/km	0-120
Roughness, BI	mm/km	2,000-14,000
Speed, S	km/h	
Cars		20 - 110
Utilities		10 - 100
Light trucks and small buses		5 - 90
Medium trucks and large buses		5 - 90
Paved roads		
Rise plus fall, RF	m/km	0-120
Roughness, BI	mm/km	2,000-9,000
Speed, S	km/h	
Car		20-140
Utilities		10-110
Light trucks and small buses		5-100
Medium trucks and large buses		5-100
Vehicle characteristics		
Power-to-weight ratio [1], PWR	Metric hp/ton	11-35
Gross vehicle weight [2], GVW	Tons	5.0-13.0

[1] For trucks and buses.
[2] For medium trucks and buses.
Source: Based on data from Hide et al., (1975).

In the above relationships lower cut-off levels have been included to avoid producing unreasonably low values of tire wear for roads smoother than those from which the survey data were collected. The cut-off levels have been set to 0.03 tires per 1,000 vehicle-km for passenger cars and utilities and to 0.01 tires per 1,000 vehicle-km per ton of vehicle weight for trucks and buses.

In using the tire wear relationships the user should observe the recommended range of the input variables shown in Table 6.3, as deviation from the range represents an extrapolation.

6.2.5 Maintenance Parts

The consumption of vehicle spare parts is expressed as a fraction of the average cost of a new vehicle per 1,000 vehicle-km of operation. The TRRL-Kenya study (Hide, et al., 1975) obtained the following relationships which relate spare parts consumption to road roughness and and vehicle age expressed in terms of cumulative kilometerage driven:

Table 6.3: Recommended range of variables for tire wear prediction: Kenya relationships

Variable	Units	Recommended range
Roughness, BI	mm/km	2,000-8,000
Gross vehicle weight GVW[1]	Tons	5.0 -26.0

Note: For trucks and buses.
Source: Adapted from Hide et al., 1975.

1. **Passenger cars and utilities**

$$PC = \begin{cases} CKM\ [-2.03 + 0.0018\ BI']\ 10^{-8} & \text{for } CKM \leq 200{,}000 \text{ km} \\ 200\ [-2.03 + 0.0018\ BI']\ 10^{-5} & \text{for } CKM > 200{,}000 \text{ km} \end{cases}$$

2. **Buses**

$$PC = \begin{cases} \sqrt{CKM}\ [-67 + 0.06\ BI']\ 10^{-8} & \text{for } CKM \leq 1{,}400{,}000 \text{ km} \\ 1.183\ [-67 + 0.06\ BI']\ 10^{-5} & \text{for } CKM > 1{,}400{,}000 \text{ km} \end{cases}$$

3. **Light and medium trucks**

$$PC = \begin{cases} CKM\ [0.48 + 0.00037\ BI]\ 10^{-8} & \text{for } CKM < 500{,}000 \text{ km} \\ 500\ [0.48 + 0.00037\ BI]\ 10^{-5} & \text{for } CKM > 500{,}000 \text{ km} \end{cases}$$

where PC = the predicted parts consumption per 1,000 vehicle-km, expressed as a fraction of the average new vehicle cost;

$$BI' = \begin{cases} BI & \text{for } BI > 1{,}500 \text{ mm/km} \\ 1500 & \text{for } BI \leq 1{,}500 \text{ mm/km} \end{cases}$$

CKM = the average cumulative kilometerage, defined as the average number of kilometers driven over the lifetime of a vehicle.

The following equation is used to compute CKM:

$$CKM = 1/2\ LIFE_o\ AKM_o$$

where $LIFE_o$ = the average vehicle service life in years, input by the user; and
AKM_o = the average number of kilometers driver per vehicle per year for the vehicle group, input by the user.

The subscript "$_o$" is used to emphasize that the values of these variables are to be provided by the user as opposed to the values calculated in terms of speeds in Section 5.3.3 (on depreciation and interest). In using these relationships the recommended range of roughness

and cumulative kilometerage compiled in Table 6.4, should be observed. As high values of average cumulative kilometerage, CKM, can lead to unrealistically high parts consumption, cut-off levels for CKM have been introduced in these relationships (i.e., 200,000 km for passenger cars and utilities, 500,000 km for trucks and 1,400,000 km for buses).

6.2.6 Maintenance Labor

The maintenance labor-hour requirements were found by TRRL (Hide et al., 1975) to be related to parts consumption and road roughness:

1. **Passenger cars and utilities**

$$LH = \begin{cases} PC\,[851 - 0.078\,BI] & BI < 6{,}000 \text{ mm/km} \\ 383\,PC & BI > 6{,}000 \text{ mm/km} \end{cases}$$

2. **Light and medium trucks**

$$LH = \begin{cases} PC\,[2975 - 0.078\,BI] & BI < 6{,}000 \text{ mm/km} \\ 2507\,PC & BI > 6{,}000 \text{ mm/km} \end{cases}$$

3. **Buses**

$$LH = \begin{cases} PC\,[2640 - 0.078\,BI] & BI < 6{,}000 \text{ mm/km} \\ 2172\,PC & BI > 6{,}000 \text{ mm/km} \end{cases}$$

where LH = the number of maintenance labor-hour per 1,000 vehicle-km;
PC = the predicted spare parts consumption, as defined earlier.

The recommended range of the input variables for predicting maintenance labor is the same as for maintenance parts prediction (Table 6.4).

Table 6.4: Recommended range of variables for maintenance parts and labor prediction: Kenya relationships

Variable	Units	Recommended range
Cumulative kilometerage, CKM	km	
Cars and utilities		0- 100,000
Light and medium trucks		0- 400,000
Buses		0-1,100,000
Roughness, BI	mm/km	2,500-7,500

Source: Adapted from Hide et al., (1975).

6.3 THE CARIBBEAN RELATIONSHIPS

The relationships described in this section were obtained from the study conducted by TRRL in Caribbean (Hide, 1982; Morosiuk and Abaynayaka, 1982). In the TRRL-Caribbean study the data were collected from vehicle operations on paved roads. Strictly speaking, therefore, the resulting relationships should be applicable only to paved roads.

6.3.1 Vehicle Speed

Using the findings by Morosiuk and Abaynayaka (1982), speeds are calculated as a function of the average rise plus fall, horizontal curvature, roughness, and carriageway width, and, for trucks, power-to-weig ratio (with the minimum limits of 20 km/h for cars and utilities, and 15 km/h for trucks.

As in the case of the Kenya relationships, the round-trip speed, S, in km/h, is defined as

$$S = \frac{2}{\frac{1}{S_u} + \frac{1}{S_d}} .$$

The expressions for S_u and S_d are as follows:

1. **Passenger cars**

$$S_u = \max\,[20;\ 67.6 - 0.078\ RF - 0.024\ C - 0.00087\ BI - 8.1 \max\,(0;\ 5 - W)]$$
$$S_d = \max\,[20;\ 67.6 - 0.067\ RF - 0.024\ C - 0.00087\ BI - 8.1 \max\,(0;\ 5 - W)]$$

2. **Utilities**

$$S_u = \max\,[20;\ 62.6 - 0.085\ RF - 0.022\ C - 0.00066\ BI - 7.0 \max\,(0;\ 5 - W)]$$
$$S_d = \max\,[20;\ 62.6 - 0.067\ RF - 0.022\ C - 0.00066\ BI - 7.0 \max\,(0;\ 5 - W)]$$

3. **Light trucks**

$$S_u = \max\,[20;\ 51.9 - 0.222\ RF - 0.017\ C - 0.00106\ BI + 0.552\ PWR - 6.2 \max\,(0;\ 5 - W)]$$

$$S_d = \max\,[20;\ 51.9 - 0.067\ RF - 0.017\ C - 0.00106\ BI + 0.552\ PWR - 6.2 \max\,(0;\ 5 - W)]$$

where the variables are as defined for the TRRL-Kenya relationships.

The user should observe the recommended range of the variables used in predicting speeds recommended compiled in Table 6.5, as deviation from this range represents an extrapolation and could produce unreasonable results.

Table 6.5: Recommended range of variables for vehicle speeds prediction: Caribbean relationships

Variable	Units	Recommended range
Road characteristics		
Rise plus fall, RF	m/km	0 - 140
Horizontal curvature, C	Degrees/km	0 - 1,200
Carriageway width, W	m	4.0 - 9.0
Roughness, BI	mm/km	1,500 -15,000
Vehicle fleet characteristics		
Power-to-weight ratio PWR	Metric hp/ton	12 - 30

Note: For light trucks.
Source: Based on data from Morosiuk and Abaynayaka (1982).

6.3.2 Fuel Consumption

The Caribbean relationships express fuel consumption in terms of the vehicle speed, average rise plus f trucks gross vehicle weight (Morosiuk and Abaynayaka, 1982).

As in the case of Kenya relationships, the predicted round-trip fuel consumption FL, in liters/1,000 km, is defined as:

$$FL = \frac{FL_u + FL_d}{2} .$$

The expressions for the FL_u and FL_d are as follows:

1. **Passenger cars**

$$FL_u = 1.16\ [24 + 969/S_u + 0.0076\ S_u^2 + 1.33\ RF\ ]$$

$$FL_d = 1.16\ [24 + 969/S_d + 0.0076\ S_d^2 - 0.63\ RF + 0.0029\ RF^2]$$

2. **Utilities**

$$FL_u = 1.16\ [72 + 949/S_u + 0.0048\ S_u^2 + 1.118\ GVW\ RF]$$

$$FL_d = 1.16\ [72 + 949/S_d + 0.0048\ S_d^2 - 1.18\ RF + 0.0057\ RF^2]$$

3. **Light trucks**

$$FL_u = 1.15\ [29 + 2219/S_u + 0.0203\ S_d^2 + 0.848\ GVW\ RF]$$

$$FL_d = 1.15\ [29 + 2219/S_d + 0.0203\ S_d^2 - 2.60\ RF - 0.0132\ RF^2]$$

where the variables are as defined for Kenya relationships. The

multiplicative factors obtained from the Brazil-UNDP Study, i.e., 1.16 for cars and utilities and 1.15 for light trucks, are also applied to the Caribbean relationships to account for the different conditions between the controlled experiments and real world operations.

Other than those variables included as explanatory variables in the above relationships, there seem to be some other factors which affect fuel consumption such as roughness, curvature and road width. However, it is assumed in the Caribbean study that the influence of these factors would be indirectly reflected through their influence on speed.

In using these fuel consumption relationships the user should observe the recommended range of the input variables shown in Table 6.6, as deviation from the range represents an extrapolation.

Table 6.6: Recommended range of variables for fuel consumption prediction Caribbean relationships

Variable	Units	Recommended range
Rise plus fall, RF [1]	m/km	0-120
Speed, S	km/h	15-80
Gross vehicle weight [2], GVW	Tons	4.0-8.5

[1] Taken to be the same as for the Kenya relationship.
[2] For light trucks.
Source: Based on data from Morosiuk and Abaynayaka (1982).

6.3.3 Tire Wear

Two relationships were derived by the TRRL-Caribbean study (Hide, 1982), one for cars and utilities, and one for light trucks. They relate the total consumption of tires per 1,000 vehicle-km to road roughness and, for light trucks, the average gross vehicle weight. These relationships are summarized below:

1. **Passenger cars and utilities**

$$TC = \begin{cases} [-60.1 + 0.0764\ BI]\ 10^{-3} & BI \geq 1{,}200\ mm/km \\ 0.03 & BI < 1{,}200\ mm/km \end{cases}$$

2. **Light trucks**

$$TC = \begin{cases} GVW\ [70.6 + 0.0135\ BI]\ 10^{-4} & BI \geq 2{,}200\ mm/km \\ 0.01\ GVW & BI < 2{,}200\ mm/km \end{cases}$$

In order to avoid producing unreasonably low values of the tire wear for smooth roads, the same cut-off levels of tire wear as those for the Kenya relationships are employed in these relationships. In using the tire wear relationships the user should observe the recommended range of the input variables shown in Table 6.7, as deviation from the range represents an extrapolation.

Table 6.7: Recommended range of variable for tire wear prediction: Caribbean relationships

Variable	Units	Recommended range
Roughness, BI	mm/km	3,000-8,000
Gross vehicle weight[1], GVW	Tons	4.0-11.0

[1] For light trucks.
Source: Based on data from Hide (1982).

6.3.4 Maintenance Parts

As in the TRRL-Kenya counterpart, the parts consumption relationships in the TRRL-Caribbean study (Hide, 1982) are expressed as a function of road roughness and vehicle cumulative kilometerage:

1. **Passenger cars and utilities**

$$PC = \begin{cases} CKM\,[-5.50 + 0.00262\,BI']\,10^{-8} & \text{for } CKM \leq 200{,}000 \text{ km} \\ 200\,[-5.50 + 0.00262\,BI']\,10^{-5} & \text{for } CKM < 200{,}000 \text{ km} \end{cases}$$

2. **Light trucks**

$$PC = \begin{cases} CKM\,[-6.54 + 0.00316\,BI' - 0.00000021\,(BI')^2]\,10^{-8} \\ \quad \text{for } CKM \leq 500{,}000 \text{ km} \\ 500\,[-6.54 + 0.00316\,BI' - 0.00000021\,(BI')^2]\,10^{-5} \\ \quad \text{for } CKM > 500{,}000 \text{ km} \end{cases}$$

where

$$BI' = \begin{cases} 2{,}500 & \text{for } BI < 2{,}500 \\ BI & \text{for } 2{,}500 \leq BI \leq 12{,}000 \\ 12{,}000 & \text{for } 12{,}000 < BI \end{cases}$$

and other variables are as defined for the Kenya relationships.

As high values of average cumulative kilometerage, CKM, can lead to unrealistically high parts consumption, cut-off levels for CKM similar to the TRRL-Kenya relationships have been introduced (i.e., 200,000 km for passenger cars and utilities, and 500,000 km for light trucks.)

In using these parts consumption relationships the user should observe the recommended range of the input variables shown in Table 6.8, as deviation from the range represents an extrapolation.

Table 6.8: Recommended range of variables for maintenance parts and labor prediction: Caribbean relationships

Variable	Units	Recommended range
Cumulative kilometerage, CKM	km	
Passenger cars		0-100,000
Utilities		0-100,000
Light trucks		0-200,000
Roughness, BI	mm/km	3,000-7,500

Source: Based on data from Hide (1982).

6.3.5 Maintenance Labor

The data obtained from the TRRL-Caribbean study (Hide, 1982) were not sufficient for estimating elaborate relationships for predicting maintenance labor requirements. Therefore, the relationships found in the TRRL-Kenya study are used, as shown in section 6.2.6 above.

6.4 THE INDIA RELATIONSHIPS

The relationships described in this section were based on the study conducted in India by the Central Road Research Institute (CRRI, 1982) and augmented by subsequent analysis by the World Bank (Chesher, 1983).

6.4.1 Vehicle Speed

Vehicle speeds are predicted as a linear function of the average rise plus fall, horizontal curvature, carriageway width, roughness, with minimum limits applied to the original relationships, i.e., 20 km/h for cars and utilities and 15 km/h for trucks and buses.

As in the case of the Kenya and Caribbean relationships, the round-trip journey speed, S, in km/h, is defined as:

$$S = \frac{2}{\frac{1}{S_u} + \frac{1}{S_u}} .$$

The expressions for S_u and S_d are as follows:

1. **Passenger cars (small and medium) and utilities:**

$$S_u = \max\left[20;\ 60.6 + 1.046\ W - 0.192\ RF - 0.0078\ C - 0.0036\ BI\right]$$
$$S_d = \max\left[20;\ 60.6 + 1.046\ W - 0.184\ RF - 0.0078\ C - 0.0036\ BI\right]$$

2. **Buses (Tata and Ashok)**

$$S_u = \max\,[15;\ 55.0 + 0.609\ W - 0.301\ RF - 0.0077\ C - 0.0022\ BI]$$
$$S_d = \max\,[15;\ 55.0 + 0.609\ W - 0.228\ RF - 0.0077\ C - 0.0022\ BI]$$

3. **Trucks (medium and heavy)**

$$S_u = \max\,[15;\ 47.3 + 1.056\ W - 0.269\ RF - 0.0099\ C - 0.0019\ BI]$$
$$S_d = \max\,[15;\ 47.3 + 1.056\ W - 0.265\ RF - 0.0099\ C - 0.0019\ BI]$$

where the variables are as defined for the Kenya relationships. The recommended range of the input variables shown in Table 6.9 should be observed, as deviation from the range could produce unreasonable results.

Table 6.9: Recommended range of variables for vehicle speeds prediction: India relationships

Variable	Units	Recommended range
Rise plus fall, RF	m/km	0-150
Horizontal curvature, C	Degrees/km	0-1,200
Carriageway width, W	m	3.5-7.0
Roughness, BI	mm/km	2,000-7,000

Source: Based on data from CRRI (1982).

6.4.2 Fuel Consumption

Fuel consumption is predicted as a function of the vehicle speed, road roughness and rise plus fall, based on the original experimentally obtained relationships adjusted to reflect actual operating conditions.

As in the case of the Kenya and Caribbean relationships, the predicted round-trip fuel consumption, FL, in liters/1,000 km, is determined as:

$$FL = \frac{FL_u + FL_d}{2}\ .$$

The expressions for FL_u and FL_d are as follows:

1. **Small passenger cars**

$$FL_u = 1.16\ [49.8 + \frac{319}{S_u} + 0.0035\ S_u^2 + 0.0019\ BI + 0.942\ RF]$$

$$FL_d = 1.16\ [49.8 + \frac{319}{S_d} + 0.0035\ S_d^2 + 0.0019\ BI - 0.677\ RF]$$

2. Medium passenger cars

$$FL_u = 1.16\left[10.3 + \frac{1676}{S_u} + 0.0133\,S_u^2 + 0.0006\,BI + 1.388\,RF\right]$$

$$FL_d = 1.16\left[10.3 + \frac{1676}{S_d} + 0.0133\,S_d^2 + 0.0006\,BI - 1.032\,RF\right]$$

3. Utilities

$$FL_u = 1.16\left[-30.8 + \frac{2260}{S_u} + 0.0242\,S_u^2 + 0.0012\,BI + 1.278\,RF\right]$$

$$FL_u = 1.16\left[-30.8 + \frac{2260}{S_d} + 0.0242\,S_d^2 + 0.0012\,BI - 0.565\,RF\right]$$

4. Medium trucks and buses (Tata)

$$FL_u = 1.15\left[85.1 + \frac{3900}{S_u} + 0.0207\,S_u^2 + 0.0012\,BI + 3.328\,RF - 4.59\,\min(PWR;\,30)\right]$$

$$FL_d = 1.15\left[85.1 + \frac{3900}{S_d} + 0.0207\,S_u^2 + 0.0012\,BI - 1.777\,RF - 4.59\,\min(PWR;\,30)\right]$$

5. Heavy trucks and buses (Ashok)

$$FL_u = 1.15\left[266.5 + \frac{2517}{S_u} + 0.0362\,S_u^2 + 0.0066\,BI + 4.265\,RF - 4.60\,\min(PWR;\,30)\right]$$

$$FL_d = 1.15\left[266.5 + \frac{2517}{S_d} + 0.0362\,S_d^2 + 0.0066\,BI + 2.737\,RF - 4.60\,\min(PWR;\,30)\right]$$

where variables are as defined for the Kenya relationships. In the equations for buses and trucks a floor value of 30 has been imposed on the power-to-weight ratio (PWR) to make sure that fuel consumption predictions do not become negative. The multiplicative factors obtained in Brazil are also used in the India relationships, i.e., 1.16 for cars and utilities, and 1.15 for trucks and buses. The recommended range of the input variables shown in Table 6.10 should be observed as deviation from the range could produce unreasonable results.

Table 6.10: Recommended range of variables for fuel consumption prediction: India relationships

Variables	Units	Recommended range
Road characteristics		
Rise plus fall, RF	m/km	0-100
Roughness, BI	mm/km	2,000-10,000
Vehicle fleet characteristics		
Speed, S	km/h	
Passenger cars and utilities		15-80
Medium trucks and buses (Tata)		10-75
Heavy trucks and buses (Ashok)		10-65
Power-to-weight ratio,[1] PWR	Metric hp/ton	7-19

[1] For trucks and buses.
Source: Based on data from CRRI (1982).

6.4.3 Tire Wear

Tire wear is predicted as a function of the road rise plus fall, horizontal curvature, width and roughness, and, in the case of buses, the cumulative kilometerage driven. The prediction employs the relationships estimated by Chesher (1983), as discussed in Appendix 6B.1, based on the data collected by CRRI (1982). As these relationships represent the life of a new tire, defined as the number of kilometers driven per tire, they have to be modified in two steps, as follows. First, they are inverted and then multiplied by the number of tires of the vehicle to represent the number of new tires consumed per 1,000 vehicle-km. This involves a bias correction for the nonlinear transformation as described in Appendix 6B. Second, a multiplicative factor equal to 0.727 obtained by CRRI (1982, Volume 2) is applied to the number of new tires consumed to account for the use of recaps; this results in the predicted number of cost-equivalent new tires per 1,000 vehicle-km:

1. **Passenger cars (small and medium) and utilities**

$$TC = [4000/(60020 - 5.86\ BI) + 0.005] \times 0.727$$

2. **Buses (Tata and Ashok)**

$$TC = [1000\ NT/\max(36100 - 0.00434\ CKM - 241\ RF - 10.54\ C - 1.126\ BI + 1044\ W;\ 2000) + 0.0022\ NT] \times 0.727$$

3. **Medium and heavy trucks**

$$TC = [1000\ NT/\max(23500 - 117.5\ RF - 8.49\ C - 0.609\ BI + 2410\ W;\ 2000) + 0.0015\ NT] \times 0.727$$

In the equation for buses and trucks, a floor value of 2,000 km has been imposed on the predicted tire life to ensure that tire wear predictions do not become negative.

The recommended range of the input variables shown in Table 6.11 should be observed, as departure from the range represents an extrapolation.

Table 6.11: Recommended range of variables for tire wear, maintenance parts and labor prediction: India relationships

Variable	Units	Recommended range
Road characteristics		
Rise plus fall, RF	m/km	
Cars and utilities		0-40
Buses		0-50
Trucks		0-60
Horizontal curvature, C	Degrees/km	
Cars and utilities		0-700
Buses		0-1,000
Trucks		0-1,200
Carriageway width, W	m	
Buses		3.5-7.5
Trucks		3.5-7.5
Vehicle fleet characteristics		
Cumulative kilometerage, CKM	km	
Cars and utilities		12,000-250,000
Buses		20,000-1,000,000
Trucks		-9,000-950,000
Gross vehicle weight, GVW	Tons	
Trucks		7.0-28.0

Source: Based on data from CRRI (1982).

6.4.4 Maintenance Parts

The relationships for predicting parts consumption were estimated by Chesher (1983) using the data collected in the India study (CRRI, 1982); these relationships are discussed in Appendix 6B.2. Since the estimation was done in log-linear form, the bias associated with the non-linear transformation must be accounted for, as described in Appendix 6B. As these relationships express parts costs in monetary units (in rupees, 1978 prices), they have to be standardized by dividing them by the average cost of a new vehicle representative of the vehicle class in question. The new vehicle prices (in rupees, 1978 prices) obtained by Chesher and Harrison

(1985) were used in the standardization, in the same manner as that employed in the Brazil parts consumption relationships. The resulting relationships express parts consumption per 1,000 vehicle-km as a function of the road roughness, and, in the case of buses and trucks, the road rise the case of buses and trucks, road roughness. The relationships employed plus fall, horizontal curvature and width, and vehicle cumulative kilometerage driven, and finally, in the case of trucks only, the gross vehicle weight:

1. **Cars (small and medium) and utilities**

$PC = PCRP_c/NVP_c$

where $PCRP_c$ = the parts cost of passenger cars, in 1978 rupees as per 1,000 vehicle-km, given by:

$PCRP_c = 42.0 \exp(0.000169\ BI)$

NVP_c = the average price of a new car in 1978 rupees, equal to Rs 64,800 which is the weighted average of the prices of the Premier Padmini and Ambassador.

2. **Buses (Tata and Ashok)**

$PC = PCRP_b/NVP_b$

where $PCRP_b$ = the parts cost of buses, in 1978 rupees per 1,000 vehicle-km, given by:

$$PCRP_b = 0.691\ (CKM)^{0.358} \exp\left(0.0000526\ BI + 0.000282\ C + 0.00675\ RF + \frac{2.00}{W}\right)$$

NVP_b = the average price of a new bus in 1978 rupees, equal to Rs234,000 which is the weighted average of the prices of the Tata and Ashok buses.

3. **Trucks (medium and heavy)**

$PC = PCRP_t/NVP_t$

where $PCRP_t$ = the parts costs of trucks, in 1978 rupees per 1,000 vehicle-km, given by:

$$PCRP_t = 0.924\ (CKM)^{0.359} \exp\left(0.0000618\ BI + 0.000686\ C + 0.000545\ RF + \frac{0.853}{W} + 0.0765\ GVW\right)$$

NVP_t = the average price of a new truck in 1978 rupees, equal to Rs 180,700 which is the weighted average of the prices of the Tata and Ashok trucks

In using the above relationship the recommended range of the input variables shown in Table 6.9 should be observed as departure from the range represents an extrapolation.

6.4.5 Maintenance Labor

Maintenance labor, expressed in the number of labor-hours per 1,000 vehicle-km, is predicted as a function of parts consumption, and in the case of buses and trucks, road roughness. The relationships employed represent a modified form of those obtained by Chesher (1983) based on the data collected in the India study (CRRI, 1982); these relationships are discussed in Appendix 6B.3. The modifications made are corrections for the biases that resulted from exponentiating the original log-linear relationships.

1. **Passenger cars (small and medium)**

$$LH = 1.799\ (PCRP_c)^{0.584}$$

2. **Utilities**

$$LH = 4.42\ (PCRP_c)^{0.445}$$

3. **Buses (Tata and Ashok)**

$$LH = 1.839\ (PCRP_b)^{0.473}\ \exp\ (0.0000426\ BI)$$

4. **Trucks (medium and heavy**

$$LH = 0.898\ (PCRP_t)^{0.654}\ \exp\ (0.0000250\ BI)$$

where LH = the number of maintenance labor hours per 1,000 vehicle-km; and

$PCRC_c$, $PCRC_b$, $PCRP_t$ = the parts costs in 1978 rupees per 1,000 vehicle-km for cars, buses and trucks, as computed previously.

In using the above relationships the recommended range of the input variables shown in Table 6.9 should be observed as departure from the range represents an extrapolation.

APPENDIX 6A

MODIFICATIONS OF TRRL-KENYA RELATIONSHIPS

6A.1 VEHICLE SPEED RELATIONSHIPS

6A.1.1 Original Equations

The speed relationships originally derived by TRRL (Hide et al., 1975) consist of separate equations for unpaved and paved roads:

Unpaved roads

1. **Passenger cars**

$$S = 84.2 - 0.210\ RS - 0.070\ F - 0.118\ C - 0.00089\ BI - 0.13\ M - 0.186\ RD$$

2. **Utilities (light goods vehicles)**

$$S = 81.2 - 0.317\ RS - 0.059\ F - 0.0966\ C - 0.00095\ BI - 0.293\ M - 0.197\ RD$$

3. **Buses**

$$S = 62.6 - 0.492\ RS - 0.0102\ F - 0.0463\ C - 0.00036\ BI - 0.163\ M - 0.0905\ RD$$

4. **Light and medium trucks**

$$S = 49.2 - 0.433\ RS + 0.00445\ F - 0.061\ C - 0.00060\ BI - 0.221\ M - 0.265\ RD + 1.10\ PWR$$

Paved roads

1. **Passenger cars**

$$S = 102.6 - 0.372\ RS - 0.0759\ F - 0.110\ C - 0.00491\ A$$

2. **Utilities**

$$S = 86.9 - 0.418\ RS - 0.0496\ F - 0.0738\ C - 0.00278\ A$$

3. **Buses**

$$S = 72.5 - 0.526\ RS + 0.0666\ F - 0.0661\ C - 0.00417\ A$$

4. Light and medium trucks

$$S = 48.0 - 0.519\ RS + 0.030\ F - 0.0581\ C - 0.00042\ A + 1.10\ PWR$$

where M = average moisture content, in percent;

RD = mean rut depth along the wheel paths, in mm; and the other variables are as defined in Section 6.2.1.

In the above equations the parameters for PWR have been modified from those of the original relationships to account for conversion from British to metric horsepower. There were also two incompatibilities between the paved and unpaved speed equations which had to be reconciled in order to make these relationships more useful:

1. The paved speed equations do not have a coefficient for the effect of road roughness. This makes predicted speeds on paved roads insensitive to road surface conditions.

2. The unpaved road speed equations do not have a coefficient for the effect of altitude, whereas the paved road speed equations do. In Kenya the average altitude of the speed test sections was about 1,300 meters above mean sea level. As altitude has a negative effect on speed in the paved road equations, when these equations were employed in predicting speeds at high altitude, it was possible for predicted speeds on unpaved roads to be greater than predicted speeds on paved roads of similar characteristics. Thus, these speed equations were inadequate for applications such as evaluating the upgrading of unpaved roads in high altitude areas.

6A.1.2 Modifications

The following paragraphs explain how roughness and altitude coefficients have been incorporated in the paved and unpaved road speed equations respectively to overcome these problems.

Roughness coefficient. A plausible reason that road roughness was found to be statistically insignificant for paved roads was that the analysis was based on a narrow range of roughness of the test sections (about 1,500 to 4,000 mm/km). On the other hand, roughness was indeed found to be statistically significant in the analysis of the unpaved road speeds in which the roughness covered a much wider range (about 3,500 to 14,000 mm/km). Therefore, the effect of roughness has been incorporated in the paved road speed relationships by extrapolating the roughness coefficients for unpaved roads, as described below.

Assume that roughness is a generic measure of road surface characteristics, meaning that paved and unpaved roads with identical roughness should have the same roughness effect on speed. Thus, the

following roughness coefficients for unpavedsurface can be transferred to paved surface situations:

Vehicle type	Roughness coefficient (km/h per mm/km)
Passenger cars	-0.00089
Utilities	-0.00095
Buses	-0.00036
Light or medium trucks	-0.00060

The transfer procedure is described as follows. For example, the original speed equation for passenger cars on paved road is given by:

$$S = 102.6 - 0.372\ RS - 0.076\ F - 0.111\ C - 0.0049\ A$$

After the coefficient transfer the equation becomes:

$$S = 102.6 - 0.372\ RS - 0.076\ F - 0.111\ C - 0.0049\ A + \Delta S - 0.00089\ BI$$

where ΔS = correction term to be added to the original constant term (102.6), so that the new equation will yield the same speed prediction when it is applied to the average roughness in the Kenya sample.

From the TRRL Report (Hide et al., 1975), this average is approximately 3,000 mm/km. Therefore, ΔS is computed as:

$$\Delta S = 0.00089 \times 3000 = 2.67\ km/h$$

The above procedure is applied to the speed equations for the remaining vehicle types. The resulting equations for all vehicle types are summarized below:

1. **Passenger cars**

$$S = 105.3 - 0.372\ RS - 0.0759\ F - 0.110\ C - 0.00089\ BI - 0.00491\ A$$

2. **Utilities**

$$S = 89.7 - 0.418\ RS - 0.0496\ F - 0.0738\ C - 0.00095\ BI - 0.00278\ A$$

3. **Buses**

$$S = 73.4 - 0.526\ RS + 0.0666\ F - 0.0661\ C - 0.00036\ BI - 0.00417\ A$$

4. **Light and medium trucks**

$$S = 49.8 - 0.519\ RS + 0.030\ F - 0.0581\ C - 0.00060\ BI - 0.00042\ A + 1.10\ PWR$$

Altitude coefficient. That altitude was not found to be a statistically significant variable for unpaved roads can possibly be attributed to the fact that except for very few sections, the unpaved test sections were at altitudes within a relatively narrow range of 1,000 - 2,000 meters. On the other hand, the altitudes for the paved test sections speed spread quite uniformly over a considerably greater range of 200 - 2,300 meters. If this argument is correct, then it would seem reasonable to use the paved road altitude coefficient in the unpaved road speed equations.

To do this, the value of 1,300 meters (above the mean sea level) is taken as the mean value for the paved road section. Using a similar procedure to that for the roughness coefficient, the following revised speed relationships are obtained for unpaved roads:

1. **Passenger cars**

$$S = 90.6 - 0.209\ RS - 0.070\ F - 0.118\ C - 0.00089\ BI - 0.135\ M - 0.186\ RD - 0.00491\ A$$

2. **Utilities**

$$S = 84.8 - 0.317\ RS - 0.059\ F - 0.0966\ C - 0.00095\ BI - 0.293\ M - 0.197\ RD - 0.00278\ A$$

3. **Buses**

$$S = 68.0 - 0.492\ RS - 0.0102\ F - 0.0463\ C - 0.00036\ BI - 0.163\ M - 0.0905\ RD - 0.00417\ A$$

4. **Light and medium trucks and buses**

$$S = 49.7 - 0.433\ RS + 0.00445\ F - 0.061\ C - 0.00060\ BI - 0.221\ M - 0.265\ RD - 0.00042\ A + 1.10\ PWR$$

In addition to the modifications described above, the relationships were further modified as described below in the HDM-III to keep the data requirement within a reasonable range:

5. **Elimination of the variables M and RD from the relationships for unpaved roads.**

In the light of the relatively insignificant contributions ofthese variables to speed prediction, they were replaced by the mean values to reduce the data requirement. The following mean values were listed in the TRRL report:

Variable	Mean value
Moisture content (%)	2.6
Rut depth (mm)	18.9

6A.2 FUEL CONSUMPTION: CHESHER NON-LINEAR RELATIONSHIPS

Using multiple regression analysis, Chesher (1977a) first estimated fuel consumption relationships for passenger cars, utilities, and trucks with three levels of loading (viz., empty, half-full, and full). The unmodified fuel consumption prediction, FL_1, in liters/1,000 km is given by:

$$FL_1 = \frac{FL_{1u} + FL_{1d}}{2}$$

where FL_{1u} and FL_{1d} are direct predictions of fuel consumption from regression equations, for the uphill and downhill segments, respectively, (in liters/1,000 km). The expressions for FL_{1u} and FL_{1d} are:

1. **Passenger car**

   $FL_{1u} = a_u + b\ RF$
   $FL_{1d} = a_d + c\ RF$

2. **Bus and trucks**
   $FL_{1u} = \exp\ (a_u + b\ RF)$
   $FL_{1d} = \exp\ (a_d + c\ RF)$

In the above equations, a_u, b_d, b and c are functions of speed, roughness and depth of loose material, given by:

$$a_u = a_0 + a_1\ S_u + a_2\ S_u^2 + a_3\ BI + a_4\ SL$$

$$a_d = a_0 + a_1\ S_d + a_2\ S_d^2 + a_3\ BI + a_4\ SL$$

$$b = b_0 + b_1\ S_u + b_2\ S_u^2;$$

$$c = c_0 + c_1\ S_d + c_2\ S_d^2$$

where S_u , S_d , BI = independent variables as defined in the text;
SL = depth of loose material, in mm; and

$$\left.\begin{array}{l} a_0\ ,\ a_1\ a_2\ ,\ a_3\ ,\ \overline{a_4} \\ b_0\ ,\ b_1\ ,\ b_2 \\ c_0\ ,\ c_1\ ,\ c_2 \end{array}\right] = \text{regression coefficients.}$$

The values of these regression coefficients for paved and unpaved roads and different types of vehicle are compiled inTable 6A-1. The three relationships obtained using the testtruck were intended for buses and light and medium trucks.

The variable SL in the expressions for a_u and a_d is replaced with the value 1.0, the average in the TRRL-Kenya study for unpaved roads. (For paved roads SL is not defined and the parameter a_4 takes on the value 0.)

In addition, further modifications are required to account for the effects f:

1. The power-to-weight ratio;
2. The paved road roughness;
3. Speed change cycles; and
4. The gross vehicle weight.

The following paragraphs describe the above modifications in some detail:

1. **Power-to-weight ratio.** The three fuel consumption relationships obtained using the test truck under empty, half and full loads correspond to fixed power-to-weight ratios of 40.6, 19.0 and 5.1 metric hp/ton, respectively. To estimate fuel consumption for buses and trucks of other power-to-weight ratios the following linear interpolation formula is used:

$$FL_2 = \begin{cases} FL_{1f} + \dfrac{PWR - 5.1}{19.0 - 5.1}\left[FL_{1h} - FL_{1f}\right] & \text{or } PWR < 19.0 \\ FL_{1h} + \dfrac{PWR - 19.0}{40.6 - 19.0}\left[FL_{1e} - FL_{1h}\right] & \text{or } PWR > 19.0 \end{cases}$$

where FL_2 = the predicted fuel consumption, computed with the effect of power-to-weight ratio accounted for, in liters/1,000 vehicle-km;

PWR = power-to-weight ratio of the vehicle in metric hp/ton; and

$FL_{1e}, FL_{1h}, FL_{1f}$ = the predicted fuel consumption, computed with therelationships for the test truck under empty,half, and full loads, respectively, in liters/1,000 vehicle-km.

For passenger cars and utilities the effect of the power-to-weight ratio was not considered significant enough to incorporate explicitly and, therefore, we have:

$$FL_2 = FL_1$$

2. **Paved road roughness.** For paved roads and all vehicle types the increase in fuel consumption due to roughness, DFL_r, is added to FL_2. For simplicity, the adjustment DFL_r isassumed to be that employed in the original TRRL-Kenya equations, i.e.,

$$DFL_r = \begin{cases} -3.3 + 0.0011\ BI & \text{for passenger cars and utilities} \\ -4.2 + 0.0014\ BI & \text{for buses and trucks} \end{cases}$$

3. **Speed change cycles.** Based on the findings of the Brazil-UNDP study, to take account of the differences between the experimental and actual operating conditions the values of fuel consumption computed for constant speeds above for both paved and unpaved roads are increased by 16 percent for passenger cars and utilities and 15 percent for buses and trucks.

4. **Gross vehicle weight.** As only one truck was employed in the field experiments the application of the fuel consumption equation obtained has to be adjusted for different gross vehicle weights. This is done by adding a correction term, DFL_{gvw}, to the fuel consumption computed for buses and trucks; DFL_{gvw} is given by:

$$DFL_{gvw} = -179.4 + 80.8\sqrt{GVW}$$

where GVW = the gross vehicle weight, in tons.

Summary of relationships. Finally, with modifications (2), (3) and (4) incorporated, the updated fuel consumption relationships can be summarized as:

Unpaved roads

$$FL = \begin{cases} 1.16\ FL_2 & \text{for passenger cars and utilities} \\ 1.15\ FL_2 - 179.4 + 80.8\sqrt{GVW} & \text{for buses and trucks} \end{cases}$$

Paved roads

$$FL = \begin{cases} 1.16\ [FL_2 - 3.3 + 0.0011\ BI] & \text{for passenger cars and utilities} \\ 1.15\ [FL_2 - 4.2 + 0.0014\ BI] - 179.4 + 80.8\sqrt{GVW} & \text{for buses and trucks} \end{cases}$$

where FL = the predicted fuel consumption, with all effects incorporated, in liters/1,000 vehicle-km.

Table 6A.1: Regression Coefficients for Kenya Non-Linear Fuel Consumption Relationships

Coefficients	a_0	a_1	a_2	a_3	a_4	b_0	b_1	b_2	c_0	c_1	c_2
Unpaved Roads											
Passenger Cars	68.61	-.02019	0.0093	0.0008902	0.04204	2.151	-0.0111	0.00003496	-0.0380	-0.01568	0.00004032
Utilities	4.580	0.001275	0.00007562	0.00001026	0.008346	0.01541	0.0001	-0.00000199	0.005673	0.0006013	0.00000479
Trucks and buses											
Empty	4.489	0.003373	0.0001143	0.00002515	0.01480	0.01989	0.0000	0.0000005	-0.03126	0.00007993	0.00000328
Half-full	5.114	0.02883	0.0004168	0.00001562	0.00345	0.01767	0.0004	0.00000502	-0.01970	0.0006130	0.00000671
Full	5.382	0.02568	0.0003415	0.00001582	0.03070	0.008792	0.0008	-0.00001069	-0.03386	-0.001797	0.00002141
Paved Roads											
Passenger cars	64.48	0.1330	0.00729	0	0	1.671	0.0047	-0.00008529	0.1072	-0.02075	0.00007821
Utilities	4.699	0.0004365	0.00009038	0	0	0.01305	0.0001	-0.00000231	0.01229	-0.0008132	0.00000592
Trucks and buses											
Empty	4.032	0.01288	0.00001888	0	0	0.02029	0.0000	-0.00000189	-0.007446	-0.001021	0.00000919
Half-full	4.922	0.01822	0.0003209	0	0	0.02071	0.0003	-0.00000654	-0.009503	-0.001433	0.00001334
Full	5.772	0.05521	0.0006378	-	0	0.001965	0.0015	-0.00001554	-0.02161	-0.001507	0.00001648

Source: Chesher (1977a).

APPENDIX 6B

MODIFICATIONS OF INDIA RELATIONSHIPS

6B.1 TIRE WEAR RELATIONSHIPS

The tire life relationships originally obtained by Chesher (1983a, 1983b) are as follows:

1. **Passenger cars and utilities**

$$TL = 60020 - 5.86\ BI$$

2. **Large buses**

$$TL = 36100 - 4.34\ \frac{CKM}{1000} - 241\ RF - 10.54\ C - 1.126\ BI + 1044\ W$$

3. **Trucks**

$$TL = 23500 - 117.5\ RF - 8.49\ C - 0.609\ BI + 2410\ W$$

where TL = the predicted tire life, in km per physically equivalent new tire; and

the other variables are as defined in the text.

The quantity used in the HDM-III for tire consumption, i.e., the number of equivalent new tires per 1,000 vehicle-km, may be obtained by dividing 1,000 NT by TL, where NT is the number of tires per vehicle. Since the inverse of TL is a non-linear transformation, the associated bias must be corrected. The Bias may be approximated by $\hat{\sigma}^2/\overline{TL}^3$, where $\hat{\sigma}^2$ and $\overline{TL}$ are unbiased estimators of the variance of the regression error and the mean of the dependent variable. The statistics reported by Chesher and Harrison (1987) needed to compute the bias correction are as follows:

Vehicle type	$\overline{TL}$	S^2_u	S^2_w
Car	30,600	8.0×10^6	28.3×10^6
Large bus	27,700	11.2×10^6	34.6×10^6
Medium or large truck	33,400	29.6×10^6	26.9×10^6

where S_u and S_w = unbiased estimators of the variances of the company and vehicle error terms, respectively.

Using these figures, for example, the bias correction term for passenger cars and utilities can be computed as follows:

$$\text{bias} = \frac{\hat{\sigma}^2}{\overline{TL}^3} = \frac{8.0 \times 10^6 + 28.3 \times 10^6}{30{,}600^3} = 1.266 \times 10^{-6}$$

Thus the modified relationship for passenger cars and utilities is given as follows:

$$\begin{aligned} TC &= 4000/(60020 - 5.86\ BI) + 4000 \times 1.266 \times 10^{-6} \\ &= 4000/(60020 - 5.86\ BI) + 0.005. \end{aligned}$$

The other relationships used in the HDM-III were obtained in a similar manner.

6B.2 PARTS CONSUMPTION RELATIONSHIPS

The parts consumption relationships originally obtained by Chesher (1983) may be expressed in HDM-III format as follows:

1. **Cars and utilities**

$$\ln\left(\frac{PCRP_c}{10}\right) = 1.264 + 0.000169\ BI$$

2. **Large buses**

$$\ln\left(\frac{PCRP_b}{10}\right) = -0.309 + 0.358 \ln\left(\frac{CKM}{1000}\right) + 0.0000526\ BI + 0.000282\ C + 0.00675\ RF + 2.00\ \frac{1}{W}$$

3. **Medium and heavy trucks**

$$\ln\left(\frac{PCRP_t}{10}\right) = 0.359 \ln\left(\frac{CKM}{1000}\right) + 0.0000618\ BI + 0.000686\ C + 0.000545\ RF + 0.853\ \frac{1}{W} + 0.0765\ GVW$$

where the variables are as defined in the text. To use these relationships they must be exponentiated and then corrected for the bias caused by the non-linear transformation.

6B.3 MAINTENANCE LABOR RELATIONSHIPS

The maintenance labor relationships originally obtained by Chesher are as follows:

1. **Passenger cars**

$$\ln (LH) = 1.896 + 0.584 \ln \frac{PCRP_a}{10}$$

2. **Utilities**

$$\ln (LH) = 2.482 + 0.445 \ln \frac{PCRP_a}{10}$$

3. **Large buses**

$$\ln (LH) = 1.652 + 0.473 \ln \frac{PCRP_b}{10} + 0.0000426\ BI$$

4. **Trucks**

$$\ln (LH) = 1.378 + 0.654 \ln \frac{PCRP_t}{10} + 0.0000250\ BI$$

where the variables are as defined in the text. To use these relationships in HDM-III they were exponentiated and then corrected for the bias caused by the non-linear transformation (Chesher, 1982).

CHAPTER 7

Benefits, Costs and Economic Analysis

The Highway Design and Maintenance Model (HDM-III) was designed to facilitate the making of economic comparisons between many alternatives, taking into account the interrelation between construction standards and maintenance operations in determining the quality of a road link and the interrelation between road quality and the operating costs of vehicles using the road. Link-alternatives may be defined by assigning to a particular link different construction and maintenance standards, traffic, and exogenous costs and benefits sets. The submodels thus far described provide the costs of construction, maintenance, and vehicle operation for any such link-alternative. Differences in these costs are part of the basis for the economic evaluation of one alternative relative to another, or of any alternative relative to the "base" case. Also involved are benefits due to increased volumes of traffic, and exogenous benefits and costs. How these elements are put together into economic analyses is the subject of this chapter.

7.1 ROAD BENEFITS AND COSTS

For every road link and link-alternative, for every year of the analysis period, the physical quantities involved in construction and in maintenance are calculated and then multiplied by their unit prices. The resulting costs for different components are identified as either capital or recurrent, as explained after the equations. From the totals of component costs in these two categories, the cost differences between alternatives for any link in a given year are calculated as follows:

Road capital cost: $\Delta CAP_{(m-n)} = CAP_m - CAP_n$

Road recurrent cost: $\Delta REC_{(m-n)} = REC_m - REC_n$

where $\Delta CAP_{(m-n)}$ = the difference in road capital cost of alternative m relative to alternative n (for a given link in a given year);

CAP_j = the total road capital cost incurred by alternative j;

$\Delta REC_{(m-n)}$ = the difference in road recurrent cost of alternative m relative to alternative n;

REC_j = the total road recurrent cost incurred by alternative j.

These differences are calculated separately for each year for each pair of link-alternatives that are to be compared.

The total road capital cost in any year comprises all the costs incurred in that year for the construction option applied to the link-alternative and all road maintenance operations that are classified as "capital." Similarly, the total road recurrent cost in the year consists of all that year's costs for maintenance operations that are classified as "recurrent." The HDM model provides a default classification for capital and recurrent maintenance operations. For paved roads the operations treated in the model as capital are pavement reconstruction, overlaying, resealing and preventive treatment. Patching and routine-miscellaneous are treated as recurrent. For unpaved roads, regravelling is the only capital maintenance operation; all other maintenance operations for unpaved roads are treated as recurrent. However, the user has the option of overriding this classification with his own.

7.2 VEHICLE BENEFITS AND COSTS

The annual economic benefits in terms of vehicle operating cost and travel time savings are calculated separately for normal and generated traffic. Normal traffic is defined as equal to the total traffic in the baseline case, and generated traffic as the traffic induced or diverted to the road by improvements relative to the baseline. The baseline case need not be either of the alternatives in the comparison. In evaluating road user benefits the model treats each vehicle group as a demand entity or market segment which is associated with its own unit price of travel. The benefits are first computed for each vehicle group and then summed over all vehicle groups to give total road user benefits. In the context of the above definitions of normal and generated traffic, the road user benefits of alternative m relative to alternative n, for a given link in a given year, are calculated using the following formulas:

(a) Vehicle operating benefits due to normal traffic:

$$\Delta VCN_{(m-n)} = \sum_i TN_i \quad [UC_{ni} - UC_{mi}]$$

(b) Vehicle operating benefits due to generated traffic[1]

$$\Delta VCG_{(m-n)} = \sum_i 1/2\ [TG_{mi} + TG_{ni}][UC_{ni} - UC_{mi}]$$

---

1 If alternative n were the baseline case, which is normally although not necessarily so, then TG_{ni} would be equal to zero, and the above equations would reduce to the traditional formulas. In other cases, for computational convenience, the division of benefits between "normal" and "generated" as calculated by the HDM varies from the usual definitions, but the total benefits are identical.

(c) Vehicle travel time benefits due to normal traffic:

$$\Delta TCN_{(m-n)} = \sum_i TN_i \, [UT_{ni} - UT_{mi}]$$

(d) Vehicle travel time benefits due to generated traffic[2]:

$$\Delta TCG_{(m-n)} = \sum_i 1/2 \, [TG_{mi} + TG_{ni}][UT_{ni} - UT_{mi}]$$

where $\Delta VCN_{(m-n)}$ = the vehicle operating benefits due to normal traffic of alternative m relative to alternative n;

TN_i = vehicle group i normal traffic in number of vehicles per year in both directions;

UC_{ji} = the average operating cost per vehicle-trip over the link for vehicle group i under alternative j (j = n or m);

$\Delta VCG_{(m-n)}$ = the vehicle operating benefits due to generated traffic of alternative m relative to alternative n;

TG_{ji} = the generated traffic of vehicle group i due to alternative j, relative to the baseline alternative (which need not be the same as either alternative n or m), in number of vehicles per year in both directions;

$\Delta TCN_{(m-n)}$ = the travel time benefits due to normal traffic of alternative m relative to alternative n;

UT_{ji} = the average travel time cost per vehicle-trip over the link for vehicle group i under alternative j (j = n or m); and

$\Delta TCG_{(m-n)}$ = the travel time benefits due to generated traffic of alternative m relative to alternative n on the given link in the given year.

The summations, $\sum_i$, are over all the vehicle groups specified by the user.

For the underlying theory and assumptions of the above formulas the user may consult, inter alia, van der Tak and Ray (1971), Walters (1968b), or Meyer and Straszheim (1971).

---

2 See footnote 1.

7.3 EXOGENOUS BENEFITS AND COSTS

The savings in vehicle operating costs and travel time endogenously calculated by the HDM model as described above normally constitute the great majority of total benefits to society of road improvements.[3] The principal exceptions are reductions in accidents, environmental impacts (e.g., noise and air pollution), and, under impeded traffic flow conditions, the increased operating costs due to congestion. Where these effects are significant, they can be estimated separately and entered as exogenous additions to the effects that are handled internally.

The exogenous benefit-cost submodel uses parameters furnished by the user to compute, for each year of the analysis, the exogenous benefits or costs and makes the results available for the model's economic analysis. Benefit and cost streams may be scheduled to begin at definite calendar years or to be triggered by the completion of a construction project. For both benefits and costs the user has the option of entering values at particular years relative to the start of a stream or defining a stream growing from an initial value at either a linear or compound rate. With the first option, values put in remain constant every year until replaced. With the other option, the user specifies an initial value and growth rate -- as a fixed increment per year or a percentage per year -- and the submodel calculates the appropriate value each year.

For each pair of alternatives being compared on a given link in a given year, the difference in net exogenous benefits is as follows:

$$\Delta EXB_{(m-n)} = EXB_m - EXC_m - EXB_n + EXC_n$$

where $\Delta EXB_{(m-n)}$ = the difference in net exogenous benefits (benefits-costs) of alternative m relative to alternative n (for a given link and year);

EXB_j = exogenous benefit for alternative j, given by the submodel for that link and year;

EXC_j = exogenous cost for alternative j, given by the submodel for that link and year.

7.4 ECONOMIC ANALYSIS AND COMPARISONS

For each pair of link-alternatives that are to be involved in a comparison, the formulas above are used to calculate, year by year, the benefits or costs of one relative to the other, combining road, vehicle, and exogenous components, both capital and recurrent. The distinction

---

[3] For theories and methodology for estimating benefits and costs that are not computed within the model, the user may consult, inter alia, Walters (1968a), and Carnemark et al. (1976).

between capital and recurrent costs is maintained solely for the purpose of computing financial budgetary constraints in cases where separate bugetary ceilings are enforced for capital and recurrent items.

Usually it is convenient to group various link-alternatives rather than examine each alternative for each link separately. This is done by grouping several links in one group and defining "group-alternatives," each of which is a set of link-alternatives that are to be treated as a unit. In other words, each group of links with one set of alternatives can be compared with the same group of links but having a different set of alternatives. The time streams of relative benefits and/or costs of appropriate pairs of link-alternatives are summed for each year to obtain the corresponding time streams for designated pairs of group-alternatives.

Thus, for any link- or group-alternative, k, in year y, the net economic benefit of alternative m relative to alternative n, $\Delta NB_{(m-n)}$, is computed as:

$$\Delta NB_{ky(m-n)} = \Delta VCN_{ky(m-n)} + \Delta VCG_{ky(m-n)} + \Delta EXB_{ky(m-n)} + \Delta TCN_{ky(m-n)}$$

$$+ \Delta TCG_{ky(m-n)} - \Delta CAP_{ky(m-n)} - \Delta REC_{ky(m-n)}$$

where $\Delta NB_{ky(m-n)}$ = the net economic benefit of alternative m relative to alternative n in year y and the variables on the right-hand side are as defined earlier, but with subscripts k and y added to indicate the link or group and year.

From the time streams of benefits or costs for the various comparative pairs of link and/or group alternatives, the model then computes net present values without discount and with up to five user-specified discount rates, as well as the internal rate of return and first-year benefits. These concepts are elaborated below.

Savings or incremental costs in financial terms and/or in foreign exchange are also computed if appropriate inputs have been given. Following all this, if the user has so specified, comparisons are repeated with changes in selected parameters, prices, etc., to ascertain the sensitivity of the results thereto.

<u>The net present value</u> of alternative m relative to alternative n, $NPV_k(m-n)$ is computed as:

$$NPV_{k(m-n)} = \sum_{y=1}^{Y} \frac{\Delta NB_{ky(m-n)}}{[1 + 0.01\ r]^{y-1}}$$

where $\Delta NB_{ky(m-n)}$ = the net economic benefit of alternative m relative to alternative n in year y;
r = the annual discount rate, in percent;
Y = the user-specified analysis period, in years.

The internal economic rate of return, denoted by r* in percent, is the discount rate at which the net present value as defined above equals zero, i.e.,

$$\sum_{y=1}^{Y} \frac{\Delta NB_{ky(m-n)}}{[1 + 0.01\ r^*]^{y-1}} = 0.$$

The above equation is solved for r* by evaluating the net present value at five percent intervals of discount rates between -95 and +500 percent, and determining the zero(s) of the equation by linear interpolation of adjacent discount rates with net present values of opposite signs. Depending on the nature of the net benefit stream, $\Delta NB_{ky(m-n)}$, it is possible to find one or multiple solutions or none at all.

The first-year benefits, which can be used as a criterion in determining the optimal timing of investments, is defined in the model as the ratio (in percent) of the net economic benefit realized in the first year after construction completion to the increase in total capital cost:

$$FY_{k(m-n)} = 100 \frac{\Delta NB_{ky^*(m-n)}}{\Delta TCC_{k(m-n)}}$$

where $FY_{k(m-n)}$ = the first-year benefits of alternative m relative to alternative n, in percent;

$\Delta NB_{ky^*(m-n)}$ = the net economic benefit of alternative m relative to alternative n in year y*, where y* is the year immediately after the last year in which the capital cost is incurred in alternative m; and

$\Delta TCC_{k(m-n)}$ = the difference in total capital cost (undiscounted) of alternative m relative to alternative n.

Since it is neither feasible nor even desirable to compare every alternative with every other one, it is put up to the user to specify one or more "studies." For each group alternative comparison the model automatically produces all the component link alternative comparisons. Up to 50 group alternative comparisons may be specified in one run, provided that the total number of group and associated link comparisons does not exceed 200.

If there are several alternatives for a given group of links, it is best to designate one of them as the "base case" and compare each of the others with it. Any one of the set may be so designated; it does not have to be a "do-nothing" or "no-project" case, although that is often convenient.

For each group-alternative included in a study, and for each of its component link-alternatives, the model will tabulate year by year the economic costs, broken down into components corresponding to the terms of the equation for $\Delta NB_{ky(m-n)}$ given above. The foreign exchange requirements will also be tabulated year by year. Each of these time series will be totaled with no discount and with up to five discount rates.

Then, for each pair of group alternatives specified, the model will tabulate the differences in component costs and benefits year by year according to the same breakdown, as well as the difference in foreign exchange requirements, and will sum them with no discount and with up to five discount rates to get the corresponding net present values. This will be done also for the pairs of link-alternatives that are included in the group-alternative pair.

The results of these comparisons and related calculations are then summarized link by link and group by group in terms of net present value for each discount rate, internal rate of return, and first year benefits.

CHAPTER 8

Expenditure Budgeting Model

In recent years, most highway expenditure programs in the World Bank borrowing countries have contained considerably more economically worthwhile projects than can be implemented within available funds. The constraints are becoming more stringent, for many of these nations have been forced to cut back their already small budgets as a result of the recent worldwide economic difficulties since the late 1970s. The problem that faces highway authorities ever more urgently is to find the best use of the limited resources. This raises a series of questions along the following line: How should available highway funding be allocated among new road construction, reconstruction, strengthening, and maintenance? What is the economically most productive allocation of investment among different functional road classes? How should a highway network rehabilitation program be scheduled so as to spread costs within the budgetary projection over time while keeping the network in reasonably serviceable condition? And so on.

Answers to questions of this type can be obtained through combined use of HDM-III and the Expenditure Budgeting Model (EBM) alluded to earlier. The HDM model is employed to predict the discounted total transport costs and consequently the net present values of different alternatives relative to a base alternative for each link in the network. The predicted economic benefits and associated costs are then transferred into the EBM, which, in turn, selects the set of alternatives that maximizes the total net present value for the entire network subject to user-specified budget constraints.

A product of collaboration between the World Bank's Transportation Department and Economic Development Institute (EDI), the Expenditure Budgeting Model[1] provides a formal mathematical optimization framework which is based on essentially the same economic principle as the HDM model, i.e., that of maximizing of net benefits. While it is simple to use, the model is still capable of analyzing all major economic tradeoffs, particularly with respect to the scale and timing of investment, as well as handling the types of project interdependencies normally found among highway investment projects (in particular, mutual exclusivity). Together with the HDM, the EBM is designed to assist highway authorities in determining short-to-medium range highway expenditure programs (5-10 years) entailing both recurrent and capital outlays on road maintenance, new construction, strengthening and rehabilitation. The methodology underlying the EBM has

1 A complete user's manual for the Expenditure Budgeting Model is given in Chapter 6, User's Manual.

been applied in part in preparing World Bank loans to developing countries.[2]

Although the present interface between the HDM and EBM models allows for only recurrent and capital expenditure constraints to be considered, the EBM as such can handle other types of resource constraints including foreign currency restriction and the maximum capacity for maintenance or construction by the local road authority or road construction industry.

A description of the EBM model is provided in the following paragraphs.[3] While the discussion is generally in the context of highway sector planning, the model can be applied to the other transport subsectors and the transport sector as a whole.

8.1 MULTIPLE-PERIOD, MULTIPLE-CONSTRAINT EXPENDITURE BUDGETING

In the course of developing the model, various methods for handling multiple-period, multiple-constraint budgeting were reviewed, including those used by highway authorities and those described in literature surveys (e.g., McKean, 1958; Beenhakker, 1976; Wilkes, 1977; Moavenzadeh et al., 1977). These methods may be classified into two broad groupings: intuitive methods and formal methods.

Intuitive methods are generally simple, easy-to-use project selection rules. For example, using what we call the cut-off rate of return method, the rule is to set the discount rate and select only the projects that have positive net present values evaluated at the discount rate (McKean, 1958). The "optimal" solution is obtained when the project selection exactly exhausts the budget availability. The rationale of the method is that the tighter the budget the greater will be the discount rate, a parallel to the notion of higher opportunity cost of funds at the budget margin. This method is implicitly based on the assumption that the budget imposed is optimal within the context of the national economy so that economic benefits from the project selection can be reinvested at the "optimal" discount rate. This assumption is internally consistent only when this discount rate turns out to equal the opportunity cost of capital in the economy at large. When budget cutbacks are severe the former can be several times the latter. In these situations, the cut-off rate of return method tends to be myopic, i.e., biased in favor of projects with relatively short lives and small capital outlays in early years.

In another intuitive method, projects are selected sequentially from the first budget period to the last in the descending order of

---

2 See Watanatada and Harral (1980) for the development of a Costa Rica 5-year road maintenance program.

3 This chapter draws heavily on an earlier paper by Watanatada and Harral (1980).

incremental benefit-cost ratios. This decision rule does not recognize the fact that, because of scale-economies such as in pavement construction, it can be more advantageous economically to implement a few relatively high-standard projects than a larger number of relatively low-standard projects in a given year. Thus, intuitive methods, although appealing because of their simplicity, must generally be used with caution.

Formal methods are based on a mathematical formulation of the problem as one of optimization under constraints. These include optimization algorithms, which guarantee a global optimum at least asymptotically (e.g., exhaustive search, integer programming, dynamic programming) and approximate algorithms, which rely on well-tested heuristic programming concepts where strict optimality is traded off against efficient use of computer resources (e.g., the Shizuo-Toyoda-Ahmed algorithm based on the concept of effective gradient, which we have used, as discussed below).

Because of the reliance on proper mathematical formulation of the objective function to be optimized against resource constraints, formal methods lend themselves to a systematic analysis of economic trade-offs -- in the same manner as trade-off analysis in the absence of budget constraints. A common example is the problem of the scale and timing of investment. Compared to the intuitive methods, formal methods have precise problem statements, and, with the availability of computers, are easy to use. Among the formal methods, the main drawback of optimization algorithms is that, although in principle they can produce a globally optimal solution, the amount of computation can be excessive for large problems. Since the value of the objective function (which is generally the sum of net present values of individual projects) can only be determined approximately, heuristic programming methods, which can handle much larger problems and yield results close to the global optimal (say, within 5 percent), are considered to be satisfactory. In using the EBM the user has a choice between an unconditionally optimal method, an asymptotically optimal method and an approximate method.

Basic definitions

We define an "investment unit" as a set of alternatives only one of which may be implemented. Different investment units are assumed to be mutually independent. Let subscript km denote a physical project which is alternative m for investment unit k. Let K be the total number of investment units considered for implementation and M_k the number of alternatives for investment unit k. In the context of HDM, k represents a link, m a link-alternative, and K the number of links in the network. A set of K projects made up of one alternative for each of the investment units is called a "program selection."

For the project km, define

NPV_{km} = the present value of net economic benefits, expressed relative to a base case, and measured in terms of discounted future consumption stream at an exogenously

determined social discount rate[4] S (or interest rate r in HDM context); and

R_{kmqt} = the (undiscounted) amount of resource of type q, q = 1 to Q, incurred by the sectoral agency in a budget period t, t = 1 to T, where Q = the total number of resource types, and T = the total number of budget periods (the duration of t may be one or more years and need not be equal for different budget periods).

In the context of HDM, two types of resources are considered: capital and recurrent expenditures, as denoted by subscripts q = 1 and 2, respectively. The amounts of these resources are given by:

Capital: $$R_{km1t} = \sum_{y \epsilon t} CAPF_{kmy}$$

Recurrent: $$R_{km2t} = \sum_{y \epsilon t} RECF_{kmy}$$

where the terms $CAPF_{kmy}$ and $RECF_{kmy}$ are, respectively, the capital and recurrent expenditures (expressed in financial terms) incurred in year y by alternative m on link k; and the summations cover all years that fall within budget period t.

Interpretation of investment units

Defined as a set of mutually exclusive alternatives, an investment unit can be interpreted in a variety of forms, as described in the examples below.[5]

1. **Independent projects.** An investment unit may represent a single alternative project independent of other projects.

2. **Scale of investment.** A road construction project in which one of three different design standards representing alternative investment levels is to be implemented immediately.

3. **Scale and timing of investment.** The above road construction project of three alternative investment levels may have two alternative construction dates (e.g., immediately and five years from now); this gives a maximum of six mutually exclusive combinations.

---

[4] The social discount rate is basically a policy parameter which must be decided at the policy level. Discussions of the social discount rate and how it should be determined are given in Lind et al. (1982).

[5] Interpretations of mutually exclusive alternatives have been made in Hirshleifer et al. (1960), Beenhakker (1976), and Juster and Pecknold (1976).

4. **Staging strategies.** Two alternatives are considered in a road construction project: the first consists of 10 units of lump sum investment in the first year; and the second consists of 6 units of lump sum investment in each of the first and seventh years, where the seventh year is for upgrading to a higher pavement standard.

5. **Recurrent expenditure.** An investment unit may involve recurrent expenditure on an existing facility such as alternative maintenance standards for a given road class.

6. **Compound projects.** An investment unit may consist of the proposed modernization of a sugar refinery as one undertaking and the proposed upgrading of the access road to the plant site as another. The total benefits of both undertakings when taken together exceed the sum of the benefits of each if taken alone. In this example we have three combinations or compound projects: modernizing the refinery alone, upgrading the access road alone, and doing both.

The notion of the investment unit provides a convenient means for organizing project analyses. Physical projects in different investment units are assumed to be independent whereas those within each investment unit are assumed to be interdependent. In principle any type of project interdependency (mutual exclusivity, joint effect on benefits and costs, etc.) can be handled. However, the number of feasible combinations of interdependent projects should not in practice become unmanageably large.

While one may argue that all projects are interdependent at least to some extent and that no project should be considered in isolation, many projects can be assumed to be independent for approximation purposes without seriously affecting project selection results. The analyst must exercise judgment to decide whether the projects should be analyzed as independent or interdependent. Also it is assumed that the analyst has sufficient understanding of the basic engineering-economic tradeoffs of investment strategies in the sector so that only a few but meaningful alternatives are carefully formulated.

Problem statement

The public expenditure budgeting problem for multiple periods can be stated as an integer programming problem of maximizing the total net present value for the sector,[6] TNVP:

$$\text{Maximize} \quad TNPV\,[X_{km}] = \sum_{k=1}^{K} \sum_{m=1}^{M_k} NPV_{km}\, X_{km}$$

[6] Other social welfare objectives such as income distribution are not addressed herein. However, it is possible to incorporate other objectives in the form of relative weights and constraints in a mathematical programming formulation (Steiner, 1969; UNIDO, 1972; Major, 1973).

over the "zero-one" decision variables X_{km}, m = 1 to M_k and k = 1 to K, where X_{km} equals 1 if alternative m of investment unit k is chosen for implementation, and equals zero otherwise; subject to:

1. The resource constraints:

$$\sum_{k=1}^{K} \sum_{m=1}^{M_k} R_{kmqt} X_{km} \leq TR_{qt}, \; q = 1, \ldots, Q; \; t = 1, \ldots, T$$

where TR_{qt} = the maximum amount of resource type q available for budget period t; and

2. The mutual exclusivity constraints:

$$\sum_{m=1}^{M_k} X_{km} \leq 1, \; k = 1, \ldots, K$$

i.e., for each investment unit k no more than one alternative can be implemented.

If we denote by M the average number of alternatives for an investment unit, the problem formulated above has KM (= K x M) zero-one variables and QT (= Q x T) resource constraints and K interdependency constraints. The parameters which define the problem size are K, M and QT. Depending on the solution method used, different problem-size parameters determine whether the method is suitable for the problem in terms of comutational effort (CE) needed.

Solution methods

Three methods for solving the above optimization problem are provided in the EBM:

1. **Total enumeration (TOTE).** This is the unconditionally optimal solution method in the EBM. It computes the total net present values of all feasible program selections and chooses the one with the largest value.

   The computational effort needed for the TOTE method may be approximately expressed as follows:

   $$CE = C_1 M^K QT$$

   where C_1 is a constant. Thus the crucial parameters are K and M. Because of the large computational effort involved, the EBM permits the TOTE method to be employed only when the number of program selections (M^K) does not exceed a preset

limit (currently equal to 2 million). Thus, the TOTE method is feasible only for problems where the number of investment units and the number of the alternatives per investment unit are relatively small.

2. **Dynamic programming (DPGM).** This is the asymptotically optimal solution method provided in EBM. It is an extensively revised version[7] of the traditional solution method outlined in Watanatada and Harral (1980). The details of the method are given in Appendix 8A. It involves dividing the feasible space into a number of neighborhoods or windows. The method is asymptotically globally optimal in the sense that as the number of windows is increased the solution approaches the global optimum.

   If we divide each resource for each budget period into an average of L intervals, the number of windows would be L^{QT} and the computational effort may be expressed approximately as:

   $$CE = C_2 \; K \; M \; L^{QT}$$

   where C_2 is a constant. Thus the crucial problem-size parameter for this method is QT. Past experience with the method indicates that L should be of the order of 100 for a near-optimal solution. Currently, there is a preset limit of 1 million on the number of windows. Thus the DPGM method may be considered for problems where the number of resource constraints, QT, is three or fewer. With the value of QT restricted to three, practical applications of the DPGM method are in general confined to problems which have one type of resource constraint ($Q = 1$), e.g., both capital and recurrent components combined, and three or fewer budget periods ($T \leq 3$). Note: At this writing, the DPGM method is implemented only on Burroughs mainframes using both ALGOL and Fortran.

   It has been found that in some pathological cases with L equal to or smaller than 20, the "optimal" solution may change depending on the ordering of the investment units.

3. **Effective gradient (AHMED).** This is the approximate solution method provided in the EBM. It is a slightly modified version of the model for highway maintenance resource allocation developed by Ahmed (1983). Ahmed used the heuristic programming concept of effective gradient introduced by Shuizuo and Toyoda (1968) in the context of investment units with single alternatives and adapted it to the case of multiple alternatives. Appendix 8B describes the algorithm in detail.

---

[7] The revision is documented in the EBM Programmer's Guide (Rich and Vurgese, 1987).

The computational effort for the AHMED method is a polynomial function of the problem-size parameters. Thus the method is not only much faster, it can handle much larger problems than the other methods. Currently, the method permits problems with up to 50 investment units, 20 alternatives, 3 resource types and 25 time periods given that the total number of resource constraints does not exceed 50 (i.e., $K \leq 50$, $M_k \leq 20$, $Q \leq 3$, $T \leq 25$ and $QT \leq 50$).

The main disadvantage of the method is that it does not guarantee a global optimum (and, further, it occasionally fails to find an extant feasible solution). However, Toyoda (1975) and Ahmed (op. cit.) report that algorithms based on effective gradient consistently yielded solutions within 4-5 percent of the global optimal. For the purposes of highway sector planning, this degree of accuracy may be considered acceptable. Provision has been made in the EBM to deal with the problem of failure to find an extant feasible solution through pre-selection of an alternative for one or more investment-units. The details are given in the User's Manual, Chapter 6.

Sensitivity analyses

Execution of the EBM model using any of the above methods automatically provides information on the sensitivity of the optimal total net present value, TNPV*, to variations in the resource budgets TR_{qt}. The sensitivity information can be expressed in terms of resource shadow prices and cut-off economic rate of return (see Section 8.3), and also in terms of the marginal project increments. The marginal project increments may be added to or subtracted from the expenditure program if the resource budgets, TR_{qt}, themselves are modified. The sensitivity information provides answers to questions of the following type: What would be the percentage loss (or gain) in the total net present value if the budgets were decreased (or increased) by, say, 20 percent? Are the existing budgets for road construction and maintenance economically optimal? If not, how much additional funding allocation will be required? How sensitive are the road design and maintenance standards to the projected highway department budgets over the next 5 or 10 years? And so on.

8.2 SINGLE-PERIOD EXPENDITURE BUDGETING: ECONOMIC INTERPRETATIONS

Section 8.1 above deals with general expenditure budgeting problems in which more than one type of resources are involved and proposed improvements can be postponed for a few or several years, such as the paving of gravel roads or, more generally, road construction projects. For these projects, expenditures can be incurred in different time periods. However, for didactic purposes, it is easier if we consider relatively simple problems involving one aggregate expenditure ($Q = 1$) and one budget period ($T = 1$). Two main types of investment situations fall into this category.

The first type covers projects that for one reason or another (e.g., engineering feasibility) will either be implemented immediately or not at all. The question of investment timing and staged-construction does not arise for these projects. The second type involves recurrent expenditure (e.g., for road maintenance) which must be provided on a continuing basis without substantial year-to-year fluctuations. For these projects it is easier to estimate benefits for an entire recurrent expenditure program (4-5 years in duration) than to isolate benefits due to spending in any individual year. For such recurrent expenditure projects the average annual cost can be used for budgeting against an average annual budget ceiling.

Incremental benefit-cost ratio method

Single-period budgeting problems with one resource type discussed above may be solved using either the total enumeration method or the dynamic programming method to yield a globally optimal solution irrespective of the functional relationship between the alternatives' net present values and resource requirements. However, when NPV_{km} is a convex function of R_{kmqt} ($q = 1$ and $t = 1$), it is possible to use a simpler solution technique alluded to earlier as the "incremental benefit-cost ratio" method, which still utilizes the concept of effective gradient (Shizuo and Toyoda, 1968; Toyoda, 1975). Since subscripts q and t are always equal to one, we now drop them from R_{kmqt} yielding R_{km}, which is interpreted as the required expenditure for alternative m of investment unit k.

For each investment unit k, let the alternatives be so arranged that R_{km} increases monotonically with m, $m = 1$ to M_k. Then, the function NPV_{km} is convex with respect to R_{km} if

$$\frac{\Delta NPV_{k,\,m-1}}{\Delta R_{k,\,m-1}} > \frac{\Delta NPV_{km}}{\Delta R_{km}} \qquad \text{for } m = 1 \text{ to } M_k$$

where
$$\Delta NPV_{km} = NPV_{km} - NPV_{k,\,m-1}$$
$$\Delta R_{km} = R_{km} - R_{k,\,m-1}$$

and NPV_{ko} and E_{ko} are defined as zero.[8] A typical shape of the relationship between NPV_{km} and R_{km} is illustrated in Figure 8.1.[9] In this figure if each straight line segment of the curve is taken as a "project increment" m, then the assumption of convexity ensures that for a given investment unit the selection of a project increment m automatically implies selection of the earlier increments (m-1, m-2, etc.)

---

[8] The possibility that ΔR_{km} equals zero is ignored for simplicity since it has no practical implications.

[9] The convexity assumption is not applicable to projects with joint positive benefits such as in the sugar refinery example, nor to projects with increasing returns to scale.

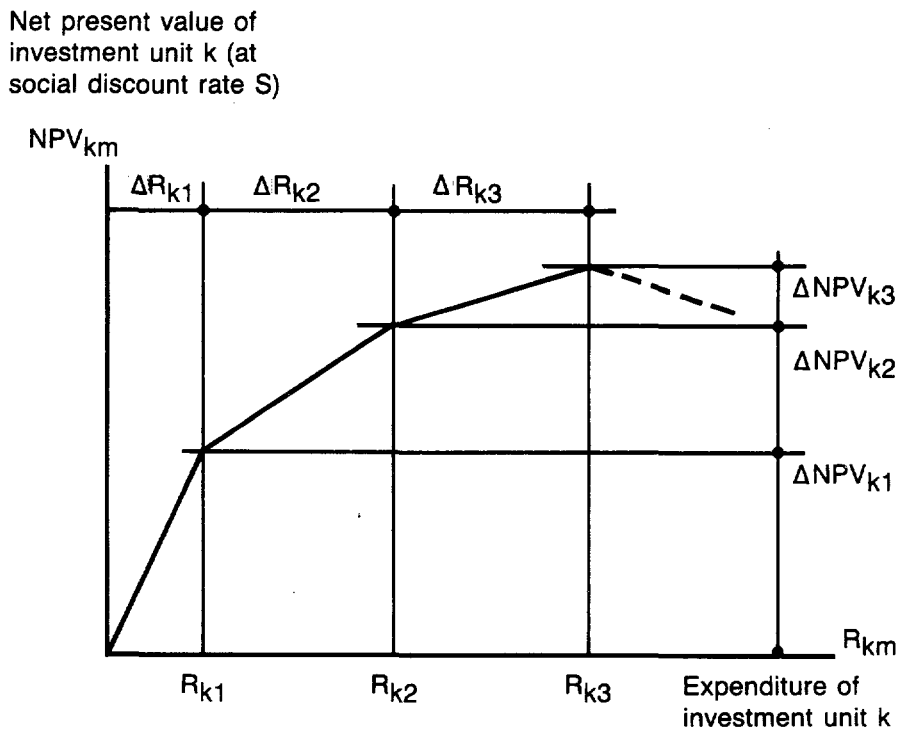
in the unit. Because of this "inclusive" property of the function NPV_{km}, the expenditure increments ΔR_{km} can be treated as if they were independent projects. For notational simplicity, a single subscript j will be made equivalent to subscripts k and m. Thus, we have ΔNPV_{km} and ΔR_{km} equivalent to ΔNPV_j and ΔR_j, respectively. For a project increment j define the "incremental benefit-cost ratio," H_j, as:

$$H_j = \frac{\Delta NPV_j}{\Delta R_j} .$$

Let N be the total number of available project increments. We now rank these project increments in the descending order of the ratio H_j, yielding a new sequence of project increments, $i = 1, 2, \ldots, N$, where[10]

$$H_i > H_{i+1} \qquad i = 1 \text{ to } N.$$

Figure 8.1: Net present value versus expenditure



[10] The possibility of a tie has been ignored for simplicity, as it has no practical implications.

The new subscript i is used instead of j to indicate randomness of the old sequence of project increments indexed by j. In the incremental B-C ratio method as many project increments are selected as the budget TR (with subscripts $q = 1$, $t = 1$ dropped from TR_{qt}) will allow by running down the sequence i, $i = 1$ to N. All project increments i, $i = 1$ to n_b are chosen where n_b is such that

$$\sum_{i=1}^{n_b} \Delta R_i \leq TR < \sum_{i=1}^{n_{b+1}} \Delta R_i .$$

Although optimality is not guaranteed the incremental B-C ratio method should be satisfactory in most cases where convexity of the net present value function applies.[11]

Economic interpretations of incremental B-C ratio at budget margin

The following interpretations are made in the context of single period expenditure budgeting but can be extended straightforwardly to multiple period budgeting.

Define the shadow premium of a budget, λ, as the incremental B-C ratio of the best project which is rejected because of the budget constraints during the given budget period. In the incremental B-C method, we have

$$\lambda = H_{n_b + 1} .$$

Then, an equivalent project selection criterion based on the "budget shadow premium" can be applied: each project increment j is selected if

$$\Delta NPV_j - \lambda \, \Delta R_j > 0 .$$

Let η denote the marginal economic rate of return obtainable in the private sector. If η is used for the social discount rate S ($S = \eta$) and no budget is imposed, we have $\lambda = 0$ and the project selection criterion reduces to the usual net present value criterion. If $S = \eta$ and the budget constraint is binding, we have $\lambda > 0$, meaning that the incremental net present value must be "penalized" by imposing a premium on the incremental investment cost ΔR_j. (Note that the incremental net present value, ΔNPV_j, already has ΔR_j factored in at face value.) Because of restricted public investment opportunities caused by the budget constraint, the cost of investment in the sector is valued at 100 λ percent higher than the cost of investment in the economy at large.

---

[11] There are cases where the solution is not optimal. For example, the total expenditure may be substantially smaller than TR so that a project increment i' where $i' > n_b + 1$ can be squeezed within the budget. But problems of indivisibility should not be serious if the investment increments are small relative to the budget.

The shadow price of incremental public expenditure budget, defined as $\lambda + 1$, can be interpreted as the present value of future benefits (for social discount rate S) that must be sacrificed for one unit of funding shortage. The future benefits can be represented by a hypothetical annual benefit stream of γ per year to perpetuity. Discounting the future benefit at social discount rate S yields

$$\lambda + 1 = \sum_{y=1}^{\infty} \frac{\gamma}{(1 + S)^y}$$

or

$$\gamma = (1 + \lambda) S.$$

The term γ may be called the "marginal (or cutoff) imputed economic rate of return" for the budget. Similarly, the "imputed economic rate of return" or "imputed ERR" for a given project increment j, $IERR_j$, is defined as

$$IERR_j = \left[\frac{\Delta NPV_j}{\Delta R_j} + 1 \right] S .$$

Using the above definitions, another equivalent criterion to the incremental B-C ratio rule can be stated on the basis of the cutoff imputed ERR. That is, we select each project increment j for which

$$IERR_j > \gamma.$$

For $S = \eta$ and where no funding constraint applies, we have $\gamma = \eta$, i.e., the cutoff imputed ERR equals the economic rate of return in the private sector. Thus the cutoff imputed ERR criterion is very similar to the more familiar internal rate of return criterion. For projects with a relatively large initial lump sum cost, uniform future consumption stream and long life, the imputed ERR is expected to be similar to the internal rate of return (IRR) computed in the regular way. For short-lived projects with a relatively small initial lump sum cost (e.g., road maintenance) the imputed ERR is expected to be considerably smaller than the regular IRR. By definition the imputed ERR is mathematically tied to the net present value which is taken as the correct measure of a project's worth. Therefore, the imputed ERR is always consistent with the net present value for project ranking purposes and does not discriminate in favor of short-lived projects as does the regular ERR.

APPENDIX 8A

DYNAMIC PROGRAMMING METHOD

One approach to solving the expenditure budgetting problem with multiple constraints presented as an integer programming problem in Section 8.2 is dynamic programming. A reformulation of the problem using dynamic programming is given first followed by a basic procedure for implementing it. Since the basic procedure was found to be relatively inefficient a new algorithm was developed to implement the dynamic programming approach. This is described next.

Dynamic Programming Formulation

For the sequence of investment units, $\ell = 1, 2, \ldots, K$, let f_k (AR_{kqt}, $t = 1$ to T and $q = 1$ to Q) denote the maximum net present value obtainable from the resources AR_{kqt} made available for the first k investment units in the sequence. Then, the available resources AR_{kqt} (where $AR_{kqt} \leq TR_{qt}$) denote state variables and subscript k stages. The function f_k is related to f_{k-1} by the following recursive relationship:

$$f_k\ (AR_{kqt},\ t = 1 \text{ to } T,\ q = 1 \text{ to } Q)$$

$$= \max \sum_{m=1}^{M_k} NPV_{km}\ X_{km} + f_{k-1}\ [AR_{kqt} - \sum_{m=1}^{M_k} R_{kmqt}\ X_{km}]$$

over the zero-one variables X_{km}, $m = 1$ to M_k; subject to the resource constraints:

$$\sum_{m=1}^{M_k} R_{kmqt}\ X_{km} \leq AR_{kqt}$$

for $t = 1$ to T and $q = 1$ to Q; and to the mutual exclusivity constraint (that no more than one alternative may be chosen):

$$\sum_{m=1}^{M_k} X_{km} < 1$$

In this dynamic programming formulation the original complex integer programming problem is reduced to K smaller subproblems each having QT state variables (AR_{kqt}, $t = 1$ to T and $q = 1$ to Q) and M_k "decision" variables (X_{km}, $m = 1$ to M_k).

Basic Dynamic Programming Algorithm

A basic procedure for obtaining a solution to this dynamic programming problem is as follows:

Step 1: Divide each total resource budget TR_{qt} into INT equal intervals. This gives L = INT + 1 discrete budget levels (or grid points) for each state variable, AR_{kqt}. This in turn gives L^{QT} combinations for period budget levels for each stage k.

Step 2: For each stage k, k = 1 to K, determine the value of f_k for each of the L^{QT} combinations of resource levels AR_{kqt}. This is accomplished by optimization over the zero-one decision variables X_{km}, m = 1 to M_k. The optimization is computationally rather trivial since it involves linear search over M_k mutually exclusive alternatives. Store each value of f_k.

Step 3: The maximum total net present value for all investment units, TNPV*, corresponding to the total resource budgets, TR_{qt}, are obtained from the recursive function at the last stage, k = K:

$$TNPV^* = f_{k=K}\ (AR_{Kqt} = TR_{qt},\ t = 1 \text{ to } T,\ q = 1 \text{ to } Q)$$

The set of projects selected corresponding to TNPV* is determined by optimization over the zero-one decision variables X_{km} for each stage in the backward order, k = K, K - 1,..., 1.

Provided that the number of grid points per budget period, L, is sufficiently large, the above algorithm yields the global optimum to the original multiple period budgeting problem. The global optimality requires no restrictive assumptions on the relationship of NPV_{km} to R_{qkmt} (Bellman, 1957).

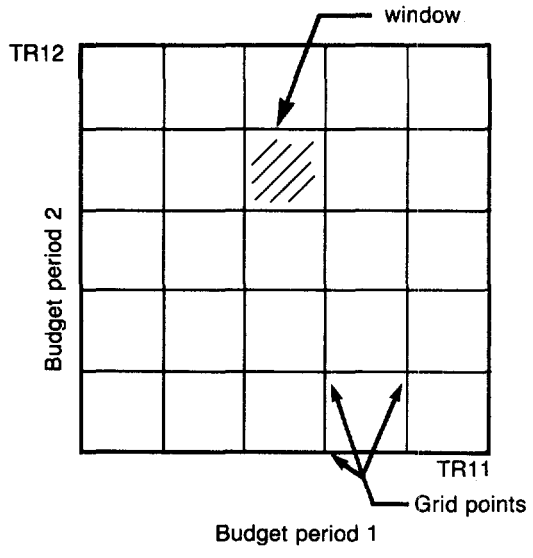
Modified Dynamic Programming Algorithm

In implementing the DP method the basic procedure above was found to be relatively inefficient. Therefore, a new algorithm was developed to take advantage of the structure of the multiple-constraint expenditure budgeting problems. Instead of storing values of f_k evaluated at grid points, the new algorithm stores values of f_k associated with actual program selections at each stage k. Feasible program selections comprise all combinations of $X_{\ell m}$, m = 1 to M_ℓ and ℓ = 1 to k that satisfy the mutual exclusivity and resource constraints. As the total number of all possible combinations, equal to

$$\prod_{\ell=1}^{k} [M_\ell + 1]$$

would normally exceed the computer memory capacity, it is necessary to store only "superior" combinations. The rule for screening out "inferior" combinations can be illustrated with the use of an example having one resource type (Q = 1) and two budget periods. Suppose that the number of intervals into which the resources TR_{11} and TR_{12} are divided equal 5. As shown in Figure 8A.1, this implies an array of grid points which form 25 squares or windows each representing a range of resources. The screening rule states that whenever two or more program selections fall into the same window only the one with the largest total net present value is kept in the storage array for stage k. The new algorithm may be summarized as follows:

Figure 8A.1 Illustrative example of one resource type and two budget periods



For each stage k. Within the total resource limits of TR_{qt}, use the alternatives for investment unit k (m = 1 to M_k plus the base alternative) to generate all feasible combinations with the ones already generated for stage k-1. (When k = 1 the number of combinations for stage 0 is defined as equal to unity which corresponds to zero expenditure and zero benefit). For each window, retain only the combination with the greatest net present value.

For the last stage K. Take the program combination with the highest net present value as the optimal solution.

In contrast to the basic procedure the new algorithm requires no backward phase as each program combination completely defines the selected alternative for each investment unit. Computational experience indicates that the total number of resource constraints of three (QT = 4) represents the maximum practical limit for the dynamic programming method to be efficient, given the number of intervals of one hundred (INT = 100) needed to maintain a reasonable degree of accuracy.

APPENDIX 8B

EFFECTIVE GRADIENT METHOD

The effective gradient method is a slight modification of the one developed by Ahmed to solve multiple-choice resource allocation problems associated with highway maintenance planning. Ahmed's method is a generalization of the approximate method proposed by Shizuo and Toyoda (1968) and Toyoda (1975) to solve binary-choice integer programming problems. Although none of these methods guarantee global optimality, these authors have reported that the solutions to the problems tested are within 4 percent of the corresponding exact solution, a degree of accuracy considered to be acceptable for highway sector planning purposes.

The effective gradient method proceeds in two stages: first, find a feasible solution based on the concept of effective gradient; and, second, search for better solutions. The computational steps are as follows:

Stage I: Find feasible solution

Step 1: For each investment unit k, consider the alternative that has the maximum net present value. Check whether all the resource constraints are satisfied. If so, go to step 10. Otherwise, go to step 2.

Step 2: For each investment unit rank the alternatives according to the ranking index, RI_{km}, defined as:

$$RI_{km} = \frac{NPV_{km}}{\sum_{q=1}^{Q} \sum_{t=1}^{T} \frac{R_{kmqt}}{TR_{qt}}} \quad m = 1, \ldots, M_k$$

Step 3: For each investment unit select the alternative that has the greatest ranking index.

Step 4: Add the resource requirements for the alternatives selected and check whether all the resource constraints are satisfied. If so, go to step 8. Otherwise, go to step 5.

Step 5: Consider the investment units in step 4 with their corresponding alternatives. Compute the effective gradients EG_k of the selected alternatives defined as:

$$EG_k = \frac{NPV_{km}}{\sum_{(q,\ t \in Q'\ T')} \sum R_{kmqt} \left(\sum_{k=1}^{K} R_{kmqt} - TR_{qt}\right)}$$

where k = 1, ..., K and Q'T' is the set of exceeded resources.

Step 6: Consider the investment unit with the smallest effective gradient and, if possible, exchange the current alternative with the next best one for the investment unit which satisfies the criterion that it be not uniformly worse in terms of requirement of exceeded resources. The next best alternative is defined in terms of the next higher ranking index RI_{km}. If all the alternatives for the investment unit are exhausted, go to step 7. Otherwise, go to step 4.

Step 7: Consider the investment unit with the next higher effective gradient. Go to step 4.

Stage II: Search for better solutions

Step 8: For each investment unit, look for an alternative which has the highest net present value other than the one currently chosen and which will be feasible if the currently chosen one is dropped out. If at least one exists, go to step 9. Otherwise, go to step 10.

Step 9: From all the investment units that have at least one better feasible alternative, select the one that gives the maximum increase in the net present value. Go to step 8.

Step 10: Stop. A final solution is obtained.

The differences between Ahmed's algorithm and the algorithm given above are located in steps 6 and 7. In step 6, Ahmed's algorithm does not check whether the exchange is uniformly worse. Further, in Ahmed's algorithm if the investment unit with the lowest effective gradient has no alternative which yields feasibility it is dropped before moving to step 7. In the above algorithm it is retained in order to increase the chance of finding a feasible program selection when one exists.

It is possible that the algorithm would fail to find an extant feasible solution in stage I. If this happens it is recommended that the user pre-select an alternative for one or more investment units and re-run the EBM. The details are given in Annex A, Chapter A6.

Glossary of Terms

A	road altitude, defined as the elevation of the road section above the mean sea level, in meters.
a_i	coefficients of fuel consumption model.
a_i	coefficients of tangential extension of the spare parts consumption model.
a_i	pavement layer strength coefficients in structural number of the road deterioration model.
AB	average floor area of bridges per unit length of road, in m^2/km.
AC	asphalt concrete
ACG	average area of site clearing and grubbing per unit length of road, in m/km.
$ACRA_a$ $ACRA_b$	total area of all cracking, comprising narrow and wide cracking, crocodile and irregular cracking of width 1 mm and greater, in percent of the total carriageway area (subscripts a, b denote after and before maintenance, respectively).
$\Delta ACRA_d$	predicted change in the area of all cracking during the analysis year due to road deterioration, in percent of the total carriageway area.
$\Delta ACRA_m$	predicted change in the area of all cracking due to maintenance, in percent of total carriageway area.
$ACRW_a$ $ACRW_b$	total area of wide cracking, comprising spalled cracks and crack widths of 3 mm or greater, in percent of the total carriageway area (subscripts a, b denote after and before maintenance, respectively).
$\Delta ACRW_d$	predicted change in the area of wide cracking during the analysis year due to road deterioration, in percent of the total carriageway area.
$\Delta ACRW_m$	predicted change in the area of wide cracking during the analysis year due to maintenance, in percent of total the carriageway area.
ADAMS	severely damaged area (i.e., damaged area for patching) at the end of the year before maintenance, in percent of the total carriageway surface area.

ADEP average annual vehicle depreciation, expressed as a fraction of the average cost of a new vehicle.

ADH average daily heavy traffic in both directions, in vehicles/day (heavy vehicles are defined as those having gross weights of 3,500 kg or more).

ADL average daily light traffic in both directions, in vehicles/day (light vehicles are defined as those having gross weights under 3,500 kg.).

ADT average daily vehicular traffic in both directions, in vehicles/day.

AF equivalent 80 kN standard axle load factor.

$AF2_k$ equivalent 80 kN standard axle load factor based on the equivalency exponent of 2.0 for vehicle group k, in ESA per vehicle.

$AF4_k$ equivalent 80 kN standard axle load factor based on the equivalency exponent of 4.0 for vehicle group k, in ESA per vehicle.

AGE1 preventive treatment age, defined as the number of years elapsed since the latest preventive treatment, reseal, overlay, pavement reconstruction or new construction.

AGE2 surfacing age, defined as the number of years elapsed since the latest reseal, overlay, pavement reconstruction or new construction.

AGE3 construction age, defined as the number of years elapsed since the latest overlay, pavement reconstruction or new construction.

AINT average annual interest on the vehicle expressed as a fraction of the average cost of a new vehicle.

AINV annual interest charge on the purchase cost of a new vehicle, in percent.

AKM predicted average number of kilometers driven per vehicle, per year.

AKM_0 baseline average number of kilometers driven per vehicle per year for the vehicle group, input by the user.

ALPC average length of regular pipe culverts, in meters.

ANBC average number of regular box culverts per unit length of road, in culverts per km.

ANBR average number of small bridges per km.

$APOT_a$ $APOT_b$	total area of potholing, in percent of the total carriageway area (subscripts a, b denote after and before maintenance, respectively).
$\Delta APOT_d$	predicted change in the total area of potholing during the analysis year due to road deterioration, in percent.
$\Delta APOT_m$	predicted change in the total area of potholing during the analysis year due to maintenance, in percent of total carriageway area.
$\Delta APOTCR_d$	predicted component change in the area of potholing during the analysis year due to cracking.
$\Delta APOT_d$	predicted component change in the equivalent area of potholing during the analysis year due to enlargement.
$\Delta APOTRV_d$	predicted component change in the area of potholing during the analysis year due to ravelling.
AR	projected frontal area of the vehicle, in m^2, which may be user-specified or take on a default value as shown in Table 5.2b.
ARV	average rectified velocity of suspension motion of the standard Opala-Maysmeter vehicle in response to road roughness, in mm/s.
ARVMAX	maximum allowable ARV, a parameter of the steady-state speed prediction model.
$ARAV_a$ $ARAV_b$	total area of ravelling, in percent of the total carriwageway area (subscripts a, b denote after and before maintenance, respectively).
$\Delta ARAV_d$	predicted change in the ravelled area during the analysis year due to road deterioration, in percent of total carriageway area.
$\Delta ARAV_m$	predicted change in the ravelled area during the analysis year due to maintenance, in percent of total carriageway area.
ΔASP	predicted area of patching performed, in percent of total carriageway surface area.
$\Delta ASPMAX$	maximum applicable patching in a year, specified by the user, in m^2/km.
$ASPS_o$	area of patching specified by the user, in m^2/km.
AXL	actual load on the axle, in kgf.
AXL_{kij}	load on axle j of a subgroup vehicle i in vehicle group k (metric tons).

B_w magnitude of the effect of road width on speed reduction, in km/h per meter of width reduction below 5 m.

BI road roughness in mm/km, as measured by the TRRL Towed Fifth Wheel Bump Integrator (Hide, et al., 1975).

C average horizontal curvature of the road, in degrees/km.

$C_{\emptyset lh}$ constant coefficient in the labor hours model.

$C_{\emptyset sp}$ constant term in the exponential relation between spare parts consumption and roughness.

$C_{\emptyset tc}$ constant term of the tire tread wear model.

C_{lhpc} parts cost exponent in the maintenance labor hours model.

C_{lhqi} roughness coefficient in the maintenance labor hours model, per QI.

C_{spqi} roughness coefficient in the exponential relation between spare parts consumption and roughness, (per QI).

C_{tctc} wear coefficient of the tire tread wear model.

CAP_{kyj} total road capital cost incurred by alternative j in analysis year y for link k.

$\Delta CAP_{ky(m-n)}$ increase in road capital cost of alternative m relative to alternative n in analysis year y.

CBR California Bearing Ratio of the subgrade at in situ conditions of moisture and density, in percent.

CD dimensionless aerodynamic drag coefficient of the vehicle, which may be user-specified or take on a default value as shown in Table 5.2b.

CF_d average circumferential force per tire on the downhill segment, in newtons.

CF_u average circumferential force per tire on the uphill road segment, in newtons.

CQ construction fault code (1 if the pavement surfacing has construction faults; 0 otherwise) specified by the user as an option or equal to the default value of zero.

CKM age of the vehicle group as proxied by the average number of kilometers the vehicles in the group have been driven, in km.

CKM' upper limit on CKM.

CMOD	resilient modulus of soil cement, in GPa (relevant only for pavements with cemented base and surface types ST or AC).
CMST	cold mix on surface treatment.
COMP	relative compaction in the base, subbase and selected subgrade layers, in percent.
CR	dimensionless coefficient of rolling resistance.
CRM	built-in standard cracking retardation time due to maintenance, in years.
CRP	delay in cracking progression due to preventive treatment.
CRPM	calibrated engine speed, in rpm.
CRX	cracking index weighted for severity of cracking at the beginning of the analysis year.
CRX	cracking index constrained such that the strength coefficient of a cracked asphalt layer is not less than that of an equivalent granular layer, min (63;MCR_a).
ΔCRX_d	predicted change in cracking index (weighted for cracking severity) due to road deterioration.
ΔCRX_m	predicted change in cracking index (weighted for cracking severity) due to maintenance.
CRT	cracking retardation time due to maintenance, in years.
CRTMAX	built-in maximum limit for CRT.
$D95_j$	maximum particle size of the material, defined as the equivalent sieve opening through which 95 percent of the material passes, in mm.
DEF	mean Benkelman beam deflection over time, in mm.
DFL_R	predicted increase in fuel consumption due to road roughness, in liters/1000 vehicle-km.
DG'	number of days between successive gradings, determined from the traffic or roughness parameter.
DGMAX	maximum allowable time interval between successive gradings, in days, specified by the user as an option or equal to the default value of 10,000 days.
DGMIN	minimum applicable time interval between successive gradings, in days, specified by the user as an option or equal to the default value of 5 days.

DRL — aggregate length of regular pipe culverts per unit length or road, in m/km.

ECR — predicted excess cracking beyond the amount that existed in the old surfacing layers at the time of the last pavement reseal, overlay or reconstruction.

EMC — equilibrium moisture content, in percent by mass.

E4RS — exponent which is a function of the surface characteristics and precipitation.

EWV — volume of earthwork per unit length of road, in m^3/km (includes cut, fill, borrow and waste materials).

EXB_{kyj} — total exogenous benefits accrued by alternative j in analysis year y.

$\Delta EXB_{ky(m-n)}$ — increase in net exogenous benefits of alternative m relative to alternative n in analysis year y.

EXC_{kyj} — total exogenous costs incurred by alternative j in analysis year y.

F_c — occurrence distribution factor for cracking initiation for the subsection (the values used in HDM-III are listed in Table 4.5).

$f(\epsilon)$ — probability density function of ϵ.

F_r — occurrence distribution factor for ravelling initiation for the subsection (the values used in HDM-III are 0.54, 0.97, and 1.49 for weak, medium, and strong subsections, respectively).

$FDAM_0$ — percentage of the damaged area specified by the user to be patched, in percent.

FL — predicted fuel consumption, in liters/1000 vehicle-km.

FL_1 — predicted fuel consumption obtained directly from the regression equations, in liters/1000 vehicle-km.

FL_{1e}, FL_{1h} — predicted fuel consumption, computed with the relationships for the test truck under empty, half and full loads, respectively, in liters/1000 vehicle-km.

FL_2 — predicted fuel consumption, computed with the effect of power-to-weight ratio accounted for, in liters/1000 vehicle-km.

$FPOT_0$ — percentage of potholing area to be patched per year, specified by the user, in percent.

FRATIO — dimensionless perceived friction ratio.

$FRATIO_0$ $FRATIO_1$	parameters of the relation between FRATIO and vehicle payload.
$FY_{k(m-n)}$	first-year benefits of alternative m relative to alternative n, in percent.
g	gravitational constant, equal to 9.81 m/s^2.
G	rise plus fall differential, in m/km.
GR	vertical gradient of a road segment, expressed as a fraction.
GRF	ground rise plus fall, in m/km.
GVW	gross vehicle weight, in (metric) tons, which may be user-specified or take on the default value shown in Table 5.2a as rated gross vehicle weight.
GVW_k	average gross vehicle weight for vehicle group k, in metric tons.
GVW_{ki}	gross vehicle weight for subgroup i in vehicle group k, in metric tons.
H	effective height of earthwork, in meters.
H_i	thickness of the ith pavement layer, in mm.
HAX_k	average number of heavy vehicle axles for vehicle group k.
HP_d HP_u	vehicle powers on the downhill and uphill road segments, in metric hp.
HPBRAKE	maximum used braking power, in metric hp.
HPDRIVE	maximum used driving power, in metric hp.
HPRATED	maximum rated power of the engine, in metric hp, which may be user-specified or take on the default value as shown in Table 6.1.
HRD_0	user-specified baseline number of hours driven per vheilce per year.
HSE	effective thickness of the surfacing layers, in mm.
$HURATIO_0$	"hourly utilization ratio" for the baseline case.
I_k	number of subgroups in vehicle group k.
IM	Thornthwaite's moisture index, in the 1955 classification where ($-100 \leq IM \leq +100$).

IRI — International Roughness Index, defining the roughness of a road profile by the $RARS_{80}$ quarter-car simulation statistic (see Sayers, Gillespie and Queiroz, 1985).

J_{ki} — number of single axles per vehicle in subgroup i of vehicle group k (a tandem axle is treated as two separate single axles).

k — ratio of asphalt concrete density at full compaction; currently 0.85 is used by the HDM.

k — exponent of vehicle age in the maintenance parts consumption model.

K_{ci} — user-specified deterioration factor for cracking initiation, multiplying time (default value = 1).

K_{cp} — user-specified deterioration factor for cracking progression, multiplying area per year (default value = 1).

K_{ge} — user-specified deterioration factor for age-roughness progression, multiplying annual fractional increment of roughness (default value = 1).

K_{gp} — user-specified deterioration factor for roughness progression, multiplying total increment of roughness (default value = 1).

K_{pp} — user-specified deterioration factor for pothole progression, multiplying each component increment of potholing (default value = 1).

K_{rp} — user-specified deterioration factor for rut depth progression, multiplying increment of rutdepth (default value = 1).

K_{vi} — user-specified deterioration factor for ravelling initiation, multiplying time (default value = 1).

KA — variable for indicating the presence of all cracking of the old surfacing layers.

KT — traffic-induced material whip-off coefficient, expressed as a function of rainfall, road geometry and material characteristics.

KW — variable for indicating the presence of wide cracking of the old surfacing layers.

L — average force per tire in the direction perpendicular to the road surface, in newtons.

LC — predicted maintenance labor cost, in Cr$/1000 vehicle-km, in January 1976 prices.

LE	axle load equivalency exponent.
LGTH	length of the roadway, in km.
LH	number of maintenance labor-hours per 1000 vehicle-km.
LIFE	predicted average vehicle service life, in years, using the varying vehicle life method.
$LIFE_o$	baseline average vehicle service life in years, input by the user.
LOAD	vehicle payload, in metric tons, which may be user-specified or take on a default value as shown in Table 5.2a.
M	average moisture content, in percent.
MG_j	slope of mean material gradation.
MGD_j	material gradation dust ratio.
MLA	predicted annual material loss, in mm/year.
MMP	annual average monthly precipitation, in meters per month.
$\Delta NB_{ky(m-n)}$	economic benefit of alternative m relative to alternative n in year y.
$\Delta NB_{ky*(m-n)}$	net economic benefit of alternative m relative to alternative n in year y*, where y* is the year immediately after the last year in which the capital cost is incurred in alternative m.
$NPV_{ky(m-n)}$	net economic benefit of alternative m relative to alternative n in year y.
NR	predicted number of retreadings per tire carcass.
NR_o	base number of retreadings per tire carcass.
NT	number of tires per vehicle.
NVP_b	average price of a new bus in 1978 rupees, equal to Rs234,000 which is a weighted average of the prices of the Tata and Ashok buses.
NVP_c	average price of a new car in 1978 rupees, equal to Rs64,800 which is the weighted average of the prices of the Premier Padmini and Ambassador.
NVP_t	average price of a new truck in 1978 rupees, equal to Rs180,700 which is the weighted average of the price of the Tata and Ashok trucks.
OC	lubricants consumption, in liters/1000 vehicle-km.

OVAC	asphalt overlay or slurry seal on asphalt concrete, or asphalt overlay on surface treatment.
P_{ki}	percentage of subgroup i vehicles in vehicle group k.
P02	amount of material passing the 2.0 mm sieve (or ASTM No. 10 sieve), in percent by mass.
P075	amount of material passing the 0.075 mm sieve (or ASTM No. 200 sieve), in percent by mass.
P425	amount of material passing the 0.425 mm sieve (or ASTM No. 40) sieve), in percent by mass.
PAX	average number of passengers per vehicle.
PC	predicted parts consumption per 1000 vehicle-km, expressed as a fraction of the average new vehicle cost.
$PCRP_b$	parts cost of buses, in 1978 rupees per 1000 vehicle-km.
$PCRP_c$	parts cost of passenger cars, in 1978 rupees as per 1000 vehicle-km.
$PCRP_t$	parts cost of trucks, in 1978 rupees per 1000 vehicle-km.
PCRA	area of all cracking before the latest reseal or overlay, in percent of the total carriageway area.
PCRW	area of wide cracking before the latest reseal or overlay, in percent of the total carriageway area.
PCRX	cracking index (weighted for severity) of the old surfacing and base layers.
PI	plasticity index of the material, in percent.
PXH	number of passenger-hours delayed per 1000 vehicle-km of travel.
PWR	vehicle power-to-weight ratio, in metric hp/ton.
QI_a QI_b	roughness of paved roads after previous maintenance. (subscript a) or before maintenance (subscript b), in QI units.
$QI_{(after)}$ $QI_{(before)}$	roughness of unpaved roads (subscript after) or before grading (subscript before), in QI units.
QI_d QI_m	predicted change in road roughness during the analysis year due to road deterioration (subscript d), or due to maintenance (subscript m), in QI units.

$QI(TG_1)$ $QI(TG_2)$	roughness of unpaved road at time TG_1 or time TG_2, respectively, in QI units.
QII_o	initial road roughness after pavement construction or reconstruction specified by the user or provided as a default value according to the surface type.
$QI_{\emptyset sp}$	transitional value of roughness in the spare parts model, in QI units.
$QIMAX_d$	estimated maximum roughness of unpaved roads surfacing material, in QI units.
$QIMAX_o$	maximum allowable roughness specified by the user, in QI units.
$QIMIN_j$	minimum roughness of unpaved road surfacing material, in QI units.
r	annual discount rate, in percent.
r*	internal economic rate of return (the discount rate at which the net present value equals zero).
RD	mean rut depth along both wheel paths, in mm.
RDM_a RDM_b	mean rut depth (along the wheel paths), in mm (subscripts a, be denoting after and before maintenance, respectively).
RDS_a RDS_b	standard deviations of rut depth (along the wheel paths), in mm (subscripts a, b denoting after and before maintenance, respectively).
ΔRDM_d ΔRDM_m	predicted change in the mean rut depth during the analysis year due to road deterioration (subscript d) or due to maintenance (subscript m), in mm.
ΔRDS_d ΔRDS_m	predicted change in the standard deviation of rut depth during the analysis year due to road deterioration (subscript d) or due to maintenance (subscript m), in mm.
$\Delta REC_{ky(m-n)}$	increase in road recurrent cost of alternative m relative to alternative n in analysis year y.
REC_{kyj}	total road recurrent cost of alternative j in analysis year y.
RF	road rise plus fall, in m/km.
RRF	ravelling retardation factor due to maintenance (dimensionless).
RH	rehabilitation indicator, value = 1 if the surface type is asphalt concrete (OVAC) or cold mix (CMST) overlay; = 0 otherwise.

RHO	mass density of air, in kg/m^3.
RL	round trip driving distance or route length, in km.
RREC	ratio of the cost of one retreading to the cost of one new tire, in percent.
RRFMAX	built-in maximum limit for RRF.
RRM	built-in standard ravelling retardation factor due to maintenance (dimensionless).
RSST	reseal on surface treatment.
RSAC	reseal on asphalt concrete.
RW	roadway width, in meters.
S	predicted speed, in km/h.
ΔS	correction term to be added to the original constant term (102.6) so that the new equation will yield the same speed prediction when it is applied to the average roughness in the Kenya sample.
S_o	baseline average vehicle speed specified by the user, in km/h.
SAS	Statistical Analysis System.
SCRA, SCRW SRAV	temporary variable defining sigmoidal function of all cracking (SCRA), wide cracking (SCRW), and ravelling (SRAV).
SL	depth of loose material, in mm.
SN	structural number of the pavement.
SN_o	user-specified structural number for the reconstructed pavement (the subscript o is used to emphasize that the values of these variables are to be provided by the user).
ΔSN_o	user-specified increment in the structural number due to the pavement reconstruction.
ΔSN_m	incremental change in structural number due to maintenance.
SNC	modified structural number.
SNC_a SNC_b	modified structural number applying for season 'a' (subscript a), or for season 'b' (subscript b).
SNCK	modified structural number adjusted for the effect of cracking.

ΔSNK	predicted reduction in the structural number due to cracking since the last pavement reseal, overlay or reconstruction (when the surfacing age, AGE2, equals zero).
SNSG	modified structural number contribution of the subgrade.
SP	superelevation of the road, in percent.
SSST	slurry seal on surface treatment.
ST	surface treatment.
TARE	vehicle tare weight, in metric tons, as given in Table 5A.1.
TC	number of cost-equivalent new tires consumed per 1000 vehicle-km.
$\Delta TCC_{k(m-n)}$	increase in total capital cost (undiscounted) of alternative m relative to alternative n.
$TCG_{ky(m-n)}$	travel time benefits due to "generated" traffic of alternative m relative to alternative n in year y.
$\Delta TCN_{ky(m-n)}$	travel time benefits due to "normal" trafic of alternative m relative to alternative n in year y.
ΔTCRA	fraction of the analysis year in which all-cracking progression applies, in years.
ΔTCRW	fraction of the analysis year in which wide cracking progression applies, in years.
TD	driving time on the section, in hours per trip.
TG_1, TG_2	time elapsed since latest grading, in days.
TG_{kyji}	vehicle group i "generated" traffic in year y due to alternative j relative to the baseline alternative (which need not be the same as either alternative n or m), in number of vehicles in both directions.
THG_0	user-specified gravel thickness after resurfacing, in mm.
ΔTHG_0	user-specified increase in the gravel thickness due to resurfacing, in mm.
TL	predicted tire life, in km per physically equivalent new tire.
TN	time spent on non-driving activities as part of the round trip tour, including loading and unloading, refueling, layovers, etc., in hours per trip.
TN_{kyi}	vehicle group i "normal" traffic in year y on link k, in number of vehicles in both directions.

ΔTRAV	fraction of the analysis year during which ravelling progression applies, in years.
TYCRA TYCRW	predicted number of years to the initiation of narrow cracking (TYCRA) or wide cracking (TYCRW), since last surfacing or resurfacing (when the surfacing age AGE2 = 0).
TYRAV	predicted number of years to ravelling initiation since the last surfacing or resurfacing (when the surfacing age AGE2 = 0).
UC_{kyji}	average operating cost per vehicle-trip over the link for vehicle group i under alternative j in year y (j = n or m).
UFC_d UFC_u	predicted unit fuel consumption for the downhill segment (subscript d), or for the uphill segment (subscript u), ml/s.
UT_{kyji}	average travel time cost per vehicle-trip over the link for vehicle group i under alternative j in year y. (j = n or m).
V	vehicle speed, in m/s.
V_d V_u	predicted steady-state speed for the downhill segment (subscript d), or for the uphill segment (subscript u), in m/s.
VBRAKE	limiting speed based on vertical gradient and braking capacity, in m/s.
$VBRAKE_d$ $VBRAKE_u$	limiting speed based on vertical gradient and braking capacity for the downhill segment (subscript d) or for the uphill segment (subscript u), in m/s.
$\Delta VCG_{ky(m-n)}$	vehicle operating benefits due to "generated" traffic of alternative m relative to alternative n in year y.
$\Delta VCN_{ky(m-n)}$	vehicle operating benefits due to "normal" traffic of alternative m relative to alternative n in years.
$\Delta VDG_{ky(m-n)}$	vehicle operating benefits due to "generated" traffic of alternative m relative to alternative n in year .
VCURVE	limiting speed determined by road curvature, in m/s.
VDRIVE	limiting speed based on vertical gradient and engine power, in m/s.
$VDRIVE_d$ $VDRIVE_u$	limiting speed based on vertical gradient and driving capacity for the downhill segment (subscript d), or for the uphill segment (subscript u), in m/s.
VEHG	traffic interval between successive gradings, in vehicles, specified by the user.

VGR — in-place volume of gravel material added to gravel resurfacing, in m^3/km.

VDESIR — desired speed, in the absence of other constraints based on psychological, economic, safety and other considerations.

$VDESIR_0$ — unmodified value of VDESIR obtained in the Brazil-UNDP study as referred to in Watanatada, et al., 1985, in km/h.

VROUGH — limiting speed based on road roughness associated ride severity, in m/s.

w — weight used for averaging over the amounts of cracking areas in the old and new surfacing layers.

W — width of the carriageway, in meters.

W' — reduction of carriageway width below 5 meters, in meters.

WS — width of one shoulder, in meters.

X — vector of the explanatory variables.

Y — user-specified analysis period, in years.

YAX — number of all vehicle axles for the analysis year, in million/lane.

YE2 — number of equivalent 80 kN standard axle loads for the analysis year based on the axle load equivalency exponent of 2.0 in million/lane (i.e. LE = 2.0).

YE4 — number of equivalent 80 kN standard axle loads for the analysis year based on the axle load equivalency exponent of 4.0, in million/lane (i.e. LE = 4.0).

α — factor for adjusting experimentally predicted fuel consumption to real-world conditions.

β σ^2 — parameter of the steady-state speed prediction model.

ϵ — error term.

θ — angle of the embankment slope, in radians.

References

Abaynayaka, S.W. "Some Techniques for Measuring Vehicle Operating Cost and Road Deterioration Parameters -- With Particular Reference to Developing Countries." TRRL supplementary report 193 UC. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1976.

Abaynayaka, S.W., G. Morosiuk, and H. Hide. "The Effect of Vehicle and Road Characteristics on Commercial Vehicle Speeds in Ethiopia." TRRL Report 271. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1975.

Ahmed, N.U. "An Analytical Decision Model for Resource Allocations in Highway Maintenance Management." Transportation Research, vol. 17A, no. 2, 1983.

American Association of State Highway and Transportation Officials. AASHTO Interim Guide for Design of Pavement Structures. Washington, D.C., 1972.

Aw, W.B. "Highway Construction Cost Model for Sector Planning in Developing Countries." S.M. thesis. Cambridge, Mass.: Massachusetts Institute of Technology, 1981.

--------, "The Generated Terrain Descripter for Use in Road Construction Cost Model." Unpublished draft working paper prepared for the World Bank. Washington, D.C.: World Bank, 1982.

Beenhakker, H.L. Handbook for the Analysis of Capital Investments. Westport, Conn.: Greenwood Press, 1976.

Bellman, R. Dynamic Programming. Princeton, N.J.: Princeton University Press, 1957.

Carnemark, C., J. Biderman, and D. Bovet. The Economic Analysis of Rural Road Projects. World Bank Staff Working Paper no. 241. Washington, D.C.: World Bank, August 1976.

Central Road Research Insititute (CRRI). Road User Cost Study in India. Final Report. 5 volumes. New Delhi, India, 1982.

Chesher, A. "India Road User Cost Survey: Pooling of Bus and Truck Data." Unpublished report prepared for the Transportation Department. Washington, D.C.: World Bank, 1983a.

----------. "Vehicle Operating Cost Equations Developed from the India Road User Cost Survey Data." Unpublished report prepared for the Transportation Department. Washington, D.C.: World Bank, 1983b.

----------. "Unbiased Prediction in Log-Linear Models." Unpublished project memorandum, ICR/AC/VLS/112/82. Washington, D.C.: World Bank, 1982.

----------. "Estimates of the Effects of Vehicle Speed and Highway Characteristics on Fuel Consumption." Unpublished report prepared for the Transportation Department. Washington, D.C.: World Bank, 1977a.

----------. "Predicting the Fuel Consumption of Medium and Heavy Goods Vehicles." Unpublished report prepared for the Transportation Department. Washington, D.C.: World Bank, 1977b.

Chesher, A. and R. Harrison. Highway Design and Maintenance Standards Model (HDM-III), Volume I, Vehicle Operating Costs: Evidence from Developing Countries. Washington, D.C.: World Bank, Transportation Department, 1987.

de Weille, Jan. Quantification of Road User Savings, World Bank Staff Occasional Paper no. 2. Washington, D.C.: World Bank, 1966.

Ellis, C.I. "Pavement Engineering in Developing Countries." TRRL Supplementary Report 537. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1979.

GEIPOT. Research on the Interrelationships between Costs of Highway Construction, Maintenance and Utilization (PICR). Final Report. 12 volumes. Brasilia, Brazil, 1982.

Harral, C.G. Preparation and Appraisal of Transport Projects. Prepared for the U.S. Department of Transportation. Washington, D.C., 1968.

Harral, C.G., P.E. Fossberg, and T. Watanatada. "Evaluating the Economic Priority of Highway Maintenance: Some Exploratory Analyses." Paper presented at the Pan African Conference on Highway Maintenance and Rehabilitation. Ghana, November, 1977.

Gynnerstedt, G. "A Generalized Road and Highway Traffic Simulation Model." Indo-Swedish Road Traffic Research Project. Stockholm, Sweden: Royal Institute of Technology, 1984.

----------. "Preliminary Model for Digital Computer Simulation of Motor Traffic on a Two-Lane Road' (in Swedish). Preliminary Report no. 30. National Swedish Road Research Institute, Traffic Department, 1966.

Gynnerstedt, G., A. Carlson, and B. Westerlund. "A Model for the Monte Carlo Simulation of Traffic Flow along Two-Lane Single Carriageway Rural Roads." Meddelande no. 43. National Swedish Road and Traffic Research Institute, Traffic Department, 1977.

Hide, H. "An Improved Data Base for Estimating Vehicle Operating Costs in Developing Countries." TRRL Supplementary Report 223 UC. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1976.

Hide, H., S.W. Abaynayaka, I. Sayer, and R.J. Wyatt. The Kenya Road Transport Cost Study: Research on Vehicle Operating Costs. TRRL Report LR 672. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1975.

Highway Research Board. The AASHO Road Test. National Research Council, Special Report 61E, Publication no. 944. Washington, D.C., 1962.

Hirshleifer, J., J.C. de Haven, and J.W. Milliman. Water Supply: Economics, Technology, and Policy. Chicago: The University of Chicago Press, 1960.

Hodges, J.W., J. Rolt, and T.E. Jones. The Kenya Road Transport Cost Study: Research on Road Deterioration. Report LR 673. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1975.

Jacobs, G.D. and I.A. Sayer. "Road Accidents in Developing Countries." TRRL Digest of Report SR 807. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1983.

Jones, T.E. "The Kenya Maintenance Study on Unpaved Roads." TRRL Laboratory Report 1111. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1984.

Juster, R.D. and W.M. Pecknold. "Improving the Process of Programming Transportation Investments." Transportation Research Record 599. Washington, D.C.: National Academy of Sciences, Transportation Research Board, 1976.

Lind, Robert C., and others. Discounting for Time and Risk in Energy Policy. Washington, D.C.: Resources for the Future, 1982.

Major, D.C. Multiobjective Water Resource Planning. American Geophysical Union Water Resources Monograph 4. Review draft, May, 1973.

Marglin, S.A. Public Investment Criteria. London, England: George Allen and Unwin, Ltd. Cambridge, Mass.: MIT Press, 1967.

Markow, M.J. and W.B. Aw. "Estimating Road Construction Costs for Sector Planning in Developing Countries." Transportation Research Record 898. Washington, D.C.: National Academy of Sciences, Transportation Research Board, 1983.

Marwah, B.R. "A Simulation Model of Two-Lane and Four-Lane Highways for Indian Traffic Conditions." Kanpur, India: Indian Institute of Technology, 1983 (draft).

----------. "Studies in Stochastic Modelling and Simulation of Mixed Vehicular Traffic." Ph.D. thesis. Kanpur, India: Indian Institute of Technology, 1976.

McKean, R.N. Efficiency in Government Through Systems Analysis. New York: John Wiley & Sons, Inc., 1958.

Meyer, J.R. and M.R. Straszheim. Pricing and Project Evaluation. Washington, D.C.: Brookings Institution, Transport Research Program, 1971.

Moavenzadeh, F., J.H. Stafford, J. Suhrbier, and J. Alexander. Highway Design Standards Study Phase I: The Model. World Bank Staff Working Paper no. 96. Washington, D.C.: World Bank, 1971.

Moavenzadeh, F., and B. Brademeyer. The Road Investment Analysis Model: General Framework for Link Evaluation. Series on Road Transport Planning Program for Developing Countries, vol. 2. Technology Adaptation Program. Cambridge, Mass.: Massachusetts Institute of Technology, 1977.

Moavenzadeh, F., F. Berger, B. Brademeyer, and R. Wyatt. The Highway Cost Model: General Framework. Technology Adaptation Program Report no. 75-4. Cambridge, Mass.: Massachusetts Institute of Technology, September 1975.

Moavenzadeh, F., E. Fernandez, and B. Brademeyer. "The Road Investment Analysis Model: General Fremework for Capital Budgeting and Programming." Series on Road Transport Planning Program for Developing Countries, vol. 4. TAP Report 77-7. Prepared for the Agency for International Development, United States Department of State, by the Technology Adaptation Program. Cambridge, Mass.: Massachusetts Institute of Technology, 1977.

Morosiuk, G. and S.W. Abaynayaka. "Vehicle Operating Costs in the Caribbean: An Experimental Study of Vehicle Performance." TRRL Report 1056. Crowthorne, England: Transport and Road Research Laboratory, 1982.

Palaniswamy, S.P. "A Generalized Simulation Model for Vehicle Behavior under Heterogeneous Traffic Conditions." Indo-Swedish Road Traffic Research Project. Indian Institute of Technology, 1983.

Palaniswamy, S.P., G. Gynnerstedt, and Y.R. Phull. "Indo-Swedish Road Traffic Simulation (ISWERTS) Model: A Generalized Traffic Simulation." Paper presented to the 64th Annual Meeting of the Transport Research Board, Washington, D.C., 1985.

Paterson, W.D.O. The Highway Design and Maintenance Standards Model (HDM-III), Volume III, Road Deterioration and Maintenance Effects: Models for Planning and Management. Washington, D.C.: Transportation Department, World Bank, 1987.

Rauhut, J.B., R.L. Lytton, and M.I. Darter. "Pavement Damage Functions for Cost Allocation." 4 volumes. Report FHWA-RD-84/017-020. Washington, D.C.: Federal Highway Administration, U.S. Department of Transportation, June, 1984.

Rich, J. and T. Underhill. A Programmer's Guide to the HDM Model Release III. Washington, D.C.: The World Bank, Transportation Department, 1987.

Rich, J. and Vurgese, S. A Programmer's Guide to the Expenditure Budgeting Model. Washington, D.C.: The World Bank, Transportation Department, 1987.

Robinson, R., H. Hide, J.W. Hodges, and J. Rolt. A Road Transport Investment Model for Developing Countries. TRRL Report LR 674. Crowthorne, England: Department of the Environment, Transport and Road Research Laboratory, 1975.

Sayers, M., T.D. Gillespie, and C.A.V. Queiroz. "International Experiment to Establish Correlation and Standard Calibration Methods for Road Roughness Measurements." Technical Report no. 45. Washington, D.C.: World Bank, 1986.

Sayers, M., T.D. Gillespie, and W.D.O. Paterson. "Guidelines for the Conduct and Calibration of Road Roughness Measurements." Technical Report no. 46. Washington, D.C.: World Bank, 1986.

Senju, S. and T. Toyoda. "An Approach to Linear Programming with 0-1 Variables." Management Science, vol. 15, no. 4. December, 1968.

Squire, L. and H.G. van der Tak. Economic Analysis of Projects. A World Bank research publication. Washington, D.C.: World Bank, 1975.

Steiner, P.O. Public Expenditure Budgeting. Washington, D.C.: Brookings Institution, 1969.

Toyoda, Y. "A Simplified Algorithm for Obtaining Approximate Solutions to Zero-One Programming Problems. Management Science, vol. 21, no. 12. August, 1975.

Tsunokawa, K. "Evaluation and Improvement of Road Construction Cost Models." Unpublished draft working paper prepared for the World Bank. Washington, D.C.: World Bank, 1983.

United Nations Industrial Development Organization (UNIDO). Guidelines for Project Evaluation. New York: United Nations, 1972.

van der Tak, H.G. and A. Ray. The Economic Benefits of Road Transport Projects. World Bank Staff Occasional Paper no. 13. Washington, D.C.: World Bank, 1971.

Visser, A.T. " An Evaluation of Unpaved Road Performance and Maintenance." Ph.D. thesis. Austin, Texas: University of Texas at Austin, Department of Civil Engineering, February, 1981.

Walters, A.A. The Economics of Road User Charges. World Bank Staff Occasional Paper no. 5. Washington, D.C.: World Bank, 1968a.

----------. "A Development Model of Transport." American Economic Review. May, 1968b.

Watanatada, T. and C.G. Harral. "Determination of Economically Balanced Highway Expenditure Programs Under Budget Constraints: Case Study of Costa Rica." Paper presented to the Regional Seminar on Road Maintenance. Asian Development Bank, Manila, 1980a.

----------. "Determination of Economically Balanced Highway Expenditure Programs under Budget Constraints, A Practical Approach." Paper presented at The World Conference on Transport Research. London, 1980.

Watanatada, T., A. Dhareshwar, and P.R.S. Rezende-Lima. The Highway Design and Maintenance Standards Model (HDM-III), Volume II, Vehicle Operating Costs: Evidence from Developing Countries. Washington, D.C.: Transportation Department, World Bank, 1987.

The World Bank

The Highway Design and Maintenance Standards Series

To provide road design and maintenance standards appropriate to the physical and economic circumstances of developing countries, the World Bank in 1969 instituted the Highway Design and Maintenance Standards Study, which developed into a major collaborative research project with leading research institutions and highway administrations in Australia, Brazil, France, India, Kenya, Sweden, the United Kingdom, and the United States. The aims of the study comprised the rigorous empirical quantification of cost tradeoffs between road construction, maintenance, and vehicle operating costs; and, as a basis for highway decisionmaking, the development of planning models incorporating total life-cycle cost simulation. Controlled experiments and extensive road user surveys were conducted to provide comprehensive data on highway conditions and vehicle operating costs in radically different economic environments on three continents.

The five volumes in the series represent the culmination of the 18-year endeavor, along with a computerized highway sector planning and investment model, currently in its third version (HDM-III). The first three volumes in the series provide theoretical foundations and statistical estimation of the underlying physical and economic relationships. The other two discuss the model and its use and are essential references for applying HDM-III.

Vehicle Operating Costs: Evidence from Developing Countries

Andrew Chesher and Robert Harrison

Presents an economic model of firms' management of vehicle fleets, which serves as a framework for the statistical analysis of vehicle operating cost data.

Vehicle Speeds and Operating Costs: Models for Road Planning and Management

Thawat Watanatada, Ashok M. Dhareshwar, and Paulo Roberto S. Rezende Lima

Presents the theory and estimation of a comprehensive set of models to predict speeds and operating costs under free flow conditions for a wide range of vehicles on medium- and low-volume roads as functions of road geometry and condition.

Road Deterioration and Maintenance Effects: Models for Planning and Management

William D. O. Paterson

Contains an extensive analysis of the physical processes, causes of deterioration, and performance prediction relationships, as well as the effectiveness of maintenance practices on unpaved and paved roads.

The Highway Design and Maintenance Standards Model
Volume 1. Description of the HDM-III Model
Volume 2. User's Manual for the HDM-III Model

Thawat Watanatada, Clell G. Harral, William D. O. Paterson, Ashok M. Dhareshwar, Anil Bhandari, and Koji Tsunokawa

Volume 1 organizes relationships described in the first three volumes, as well as a road construction submodel, into interacting sets of costs related to construction, maintenance, and road use. Volume 2 provides guidance on the use of this model—including input data forms, inference ranges, and default values—and gives numerical examples.

ISBN 0-8018-3091-7
ISBN 0-8018-3668-9 (5-volume set)