ARIZONA DEPARTMENT OF TRANSPORTATION

REPORT NUMBER: AZ-SP-9702

EXPERT PROJECT RECOMMENDATION PROCEDURE FOR ADOT'S PAVEMENT MANAGEMENT SYSTEM

Special Report

Prepared by:
Gerardo W. Flintsch
Arizona Transportation Research Center
Arizona Department of Transportation
1130 North 22nd Avenue
Phoenix, AZ 85009

John P. Zaniewski
Arizona State University
Civil Engineering Department
Tempe, AZ 85287

James Delton
Arizona Department of Transportation
1221 North 21st Avenue
Phoenix, AZ 85009

August 1997

Prepared for:
Arizona Department of Transportation
206 South 17th Avenue
Phoenix, Arizona 85007
in cooperation with
U.S. Department of Transportation
Federal Highway Administration

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-AZ SP-9702		

4. Title and Subtitle	5. Report Date
EXPERT PROJECT RECOMMENDATION PROCEDURE FOR ADOT'S PAVEMENT MANAGEMENT SYSTEM	August, 1997
	6. Performing Organization Code
7. Author	8. Performing Organization Report No.
Gerardo W. Flintsch, John P. Zaniewski and James Delton	
9. Performing Organization Name and Address	10. Work Unit No.
ARIZONA TRANSPORTATION RESEARCH CENTER 1130 N. 22ND AVENUE PHOENIX, AZ 85009 - 3716	11. Contract or Grant No.
12. Sponsoring Agency Name and Address	13. Type of Report & Period Covered
ARIZONA DEPARTMENT OF TRANSPORTATION 206 S. 17TH AVENUE PHOENIX, ARIZONA 85007	FINAL, 1993-97
	14. Sponsoring Agency Code

15. Supplementary Notes

Prepared in cooperation with the US. Department of Transportation, Federal Highway Administration

16. Abstract

The Arizona Department of Transportation, ADOT, uses a network level pavement management system to determine budget requirements for their annual pavement preservation program. While this is a valuable tool for preservation programming, it does not assist the engineers with the selection of projects and rehabilitation treatments. The research documented in this paper was designed to enhance the capability of ADOT's pavement management system to include project selection.

An automatic project recommendation procedure was developed and implemented in a user friendly, modular computer program. This automatic system is expected to considerably reduce the effort required to develop preservation program. It should improve the consistency of the decision process. . The final project selection will be still conducted manually by the pavement management and district engineers.

The analysis starts with a section delineation procedure that processes the milepost data into uniform roadway sections and computes the average characteristics for each section. It then computes the remaining service life of each uniform section using lineal performance equations and trigger points defined for each condition indicator. An artificial neural network simulator is used for screening and recommending roadway sections. The trained artificial neural network prepares a list of candidate sections, using the criteria learned from past selections and the current condition of all pavement sections in the network.

This preliminary list of candidate sections is further analyzed by an "expert" project recommendation procedure that recommends a preservation treatment. It also assigns a priority rating to each section in the list, and sorts the projects by priority. Funding is assigned to the highest priority sections within each roadway group until the budget recommendation provided by the network optimization process is reached.

17. Key Words		18. Distribution Statement		23. Registrant's Seal
Project Selection, Pavement Management, Section Delineation, Performance Prediction, Artificial Neural Networks.		Document is available to the US. public through the National Technical Information Service, Springfield, Virginia, 22161		Registered Professional Engineer (Civil) CERTIFICATE NO. 10720 JAMES PHILIP DELTON DATE SIGNED 7 AUG 97 Arizona U. S. A.
19. Security Classification	20. Security Classification	21. No. of Pages	22. Price	
Unclassified	Unclassified			

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

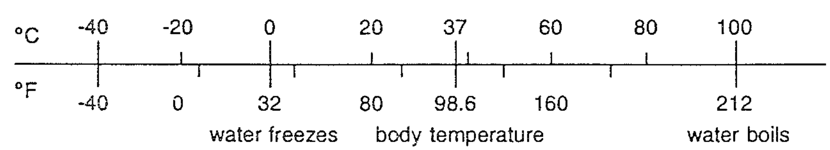
Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
ml	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	millimeters squared	mm^2
ft^2	square feet	0.093	meters squared	m^2
yd^2	square yards	0.836	meters squared	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	kilometers squared	km^2
		VOLUME		
ft oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	meters cubed	m^3
yd^3	cubic yards	0.765	meters cubed	m^3

NOTE: Volumes greater than 1000 L shall be shown in m^3.

Symbol	When You Know	Multiply By	To Find	Symbol
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg

TEMPERATURE (exact)

Symbol	When You Know	Do The Following	To Find	Symbol
°F	Fahrenheit temperature	°F - 32 ÷ 1.8	Celcius temperature	°C



APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	millimeters squared	0.0016	square inches	in^2
m^2	meters squared	10.764	square feet	ft^2
m^2	meters squared	1.19	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	kilometers squared	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	meters cubed	35.315	cubic feet	ft^3
m^3	meters cubed	1.31	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams	1.102	short tons (2000 lb)	T

TEMPERATURE (exact)

Symbol	When You Know	Do The Following	To Find	Symbol
°C	Celcius temperature	°C x 1.8 + 32	Fahrenheit temperature	°F

METER: a little longer than a yard (about 1.1 yards)
LITER: a little larger than a quart (about 1.06 quarts)
GRAM: a little more than the weight of a paper clip
MILLIMETER: diameter of a paper clip wire
KILOMETER: somewhat further than 1/2 mile (about 0.6 mile)

*SI is the symbol for the International System of Measurement

TABLE OF CONTENTS

Page

Page

Page

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

1. INTRODUCTION

A primary responsibility of federal, state and local highway agencies is to provide a network of serviceable and safe pavements at the minimum cost (AASHTO, 1990). However, at every level of government, the availability of funds is generally lower than required to maintain the roadway network at an ideal serviceability level. Therefore, finding the optimum distribution of highway funds, that results in an acceptable condition of the network, is an important goal of most governmental agencies. A Pavement Management System (PMS) can be a useful tool to help decision-makers to make the best use of the available resource by finding the optimum distribution of funds.

The Federal Highway Administration (FHWA) defines a PMS as, "a set of tools or methods that can assist decision-makers in finding cost-effective strategies for providing and maintaining pavements in a serviceable condition" (FHWA, 1989). The Interim Final Rule for Management and Monitoring Systems issued by the FHWA and the Federal Transit Administration (FTA) as a requirement of the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 provides a slightly more precise definition: "a PMS means a systematic process that provides, analyses, and summarizes pavement information for use in selecting and implementing cost-effective pavement construction, rehabilitation, and maintenance programs."

Pavement management encompasses all the activities involved in the planning, design, construction, maintenance, and rehabilitation of pavements. PMS provides a systematic methodology to improve the efficiency, expand the scope and ensure consistency of the decisions made at the different management levels within the same organization (Hudson et al., 1979).

To accomplish these objectives, a PMS must provide cost-effective decision in response to the following questions (AASHTO, 1990):

1. Where are treatments needed?
2. What treatment is the most cost-effective?
3. When is the best time (pavement condition) to program a treatment?

Two basic levels of pavement management can be considered (Hudson et al., 1979):

a. *Network management level*, where the administrative decisions that affect programs and policies for road networks are made. These decisions require general information, but handle the entire network of pavements under consideration.

b. *Project management level*, where decisions are made essentially in terms of technical management for specific projects. It involves technical concerns related to a particular road section and, therefore, requires more detailed information of this section than network level management.

Haas et al. (1994) further divides the network level into the *project selection level* and the *program level*. The project selection level involves the identification of what projects should be carried out each year of the programming period. At the program level, budgets are established and general allocations are made over an entire network.

The three-level concept is illustrated in Figure 1.1. Project level deals with specific technical concerns regarding individual sections. Therefore, it requires the most detailed information and more complex prediction models. A typical output would be a set of design strategies that minimize total life-cycle cost. The project selection level deals with decisions on funding for projects or group of projects. The pavement performance models used are less complex and require less detailed data.

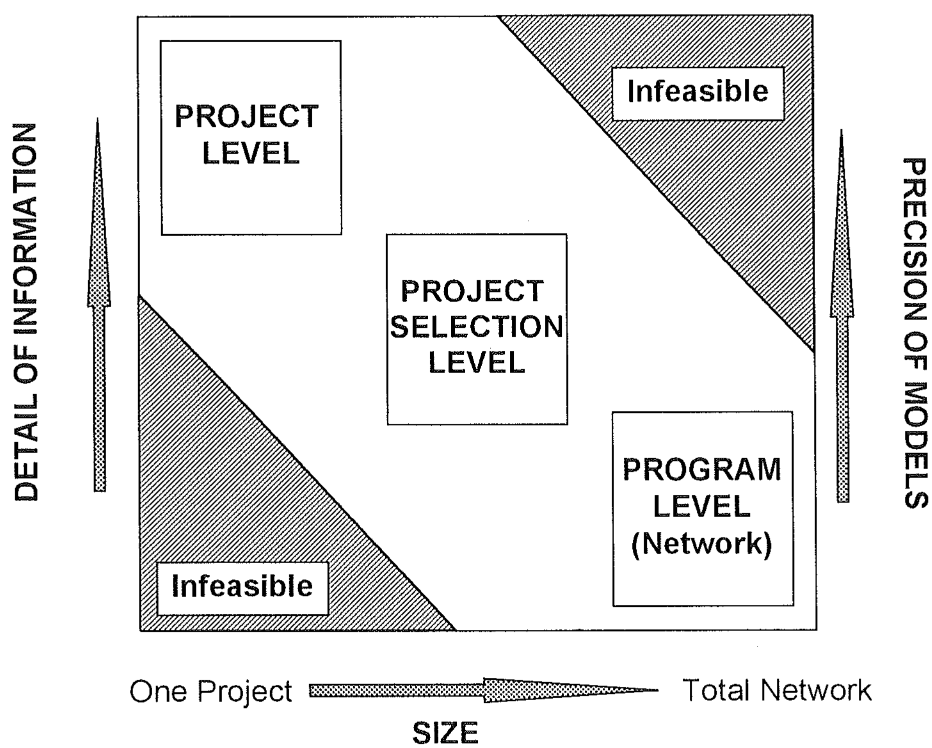


Figure 1.1. Information detail and complexity of models for a three level PMS (Haas et al. 1994)

The program level involves general budget allocation decisions for an entire network, requiring general information. The models for program level pavement management should be designed to optimize the use of funds allocated for rehabilitation and maintenance. These three levels of pavement management can be found in a number of agencies. However, it is common that the terminology overlap. For example, the expression "project level" is sometimes used when "project selection level" actually may be meant, and in other occasions the term "network level" is used when referring to the "program level" (Haas et al., 1994).

1.1. STRUCTURE OF A PAVEMENT MANAGEMENT SYSTEM

The 1990 AASHTO Guidelines for Pavement Management Systems divides a PMS in three interacting modules or "blocks." These three basic blocks are (AASHTO, 1990):

1. **Database**, containing, the inventory and condition data required for PMS analysis.
2. **Analysis Method**, to generate the products necessary for decision-making.
3. **Feedback** process, to improve the reliability of the analysis.

The degree of sophistication or completeness of the system is determined basically from the analysis method used. The AASHTO Guidelines for PMS suggest three levels:

1. **Pavement Condition Analysis**. A first level PMS requires a minimum of technology and has no prediction capabilities. The needs of the network are estimated based on the current condition. The resulting projects can be prioritized if the available money does not cover the entire set of pavements needing preservation treatments.
2. **Priority Assessment Models**. A second level PMS includes prediction models that estimate the pavement condition over time and the effect of the maintenance alternatives. This level of PMS allows to use of more sophisticated prioritization assessment methods, i.e., life-cycle cost analysis, and to estimate the effect of budget fluctuations on the condition of the pavements in the network. Using a "bottom-up" approach, optimal maintenance and rehabilitation (M&R) strategies are determined for each section in the network. The resulting projects are sorted by priority to prepare a prioritized list of projects, and a work program is prepared based on this list.
3. **Optimization Methods**. The third level of PMS analysis is defined in terms of the ability to perform network optimization. This technique evaluates the entire network over a long period of time, in order to identify a set of M&R strategies that will maximize the total network

benefits (or performance) or minimize the total cost subject to network-level constrains, such as maximum available budget and minimum desirable performance standard. This implies a "top-down" approach, in which optimal network strategies are first determined and specific treatments are then identified, considering site specific conditions and administrative policies.

1.2. OVERVIEW OF ADOT'S PAVEMENT MANAGEMENT SYSTEM

ADOT has operated a network-level PMS since 1980. This Network Optimization System (NOS) determines the optimum work program and budget for ADOT's pavement network. The system was revised and adjusted in 1993 (Wang et al., 1993). Network-level pavement condition data collection activities started in 1972 (Way and Eisenberg, 1980).

1.2.1. Antecedents

In 1976 ADOT developed, through Woodward-Clyde Consultants, WCC, a project-by-project selection framework, that provided a rationale for selecting the optimum initial structural design and maintenance strategy for new pavements and the optimum maintenance strategy for existing pavements (Finn et al., 1976). The study included the identification of relevant pavement attributes and pavement maintenance alternatives, the development of models to predict pavement performance and cost, and the establishment of appropriate "tradeoffs" or utility functions for pavement attributes to be included for optimization by the PMS. The main pavement attributes identified for the PMS were: skid number, roughness, dollar cost for routine and major maintenance, and time delays associated with major maintenance. The maintenance alternatives were divided in two groups, routine maintenance and major maintenance. Prediction models for roughness and skid number were defined subjectively. The factors used to predict roughness were traffic, deflection, environment, age and thickness. The model to predict skid resistance included aggregate type and environment only. The economic analysis considered only agency cost, such as new construction, routine and major maintenance costs, and salvage value. Variations in user cost due to the different maintenance alternatives were not considered. Finally, it was recommended that the project selection should be based on a utility function that summarized both history and uncertainty of the various attributes.

During the WCC research a computer program, SOMSAC, was developed and a parametric analysis using hypothetical pavement sections was conducted to find the sensitivity of the best maintenance strategy and its associated cost to changes in the input parameters. The main results of the analysis were (Finn et al., 1976):

1. Region has a significant influence on the best maintenance strategy. The state was divided into three regions based on elevation and soil type.
2. The deflection of a pavement also was a significant indicator of performance and useful for selection of the best alternative. Higher deflections imply weaker pavements requiring more expensive maintenance treatments.
3. The choice of the utility function parameters and weights moderately influenced the strategy selected.
4. The limiting value of roughness index has a significant effect.
5. Whether or not cracking is present did not have a significant effect on the best alternative selected.

The concepts described by Finn et al. (1976) were used intermittently in the past, but are not being currently practiced. The study recommended a second and third phase for the PMS development. The second phase consisted in the implementation of the methodology defined. The third phase was the development of a framework for network optimization.

1.2.2. ADOT PMS Database

ADOT maintains a pavement database stored in several microcomputer database files. The main PMS databases include the following distresses:

- Roughness,
- Cracking,
- Rutting,
- Flushing,
- Patching, and
- Skid resistance.

The main PMS database also includes information about: R-value of the subgrade, Structural Number (SN), maintenance cost, pavement lifts (only last project), and traffic. The data format of the main Pavement Management database is shown in Figure 1.2.

Separate databases include information about projects, containing treatments performed and dates at which each treatment was applied, and pavement deflection measurements (partial).

FIELD NAME	*A*	*RTNO*	*D*	*MP*	*L*	*PW*	*LS*	*RS*	*DIST*	*AREA*	*REG*	*ADT*	*ADL*	*GF*	->
	I	8	E	1	2	24	5	10	Y	7	0.4	11915	1247	4.3	->
	I	8	E	2	2	24	4	10	Y	7	0.4	11915	1247	4.3	->
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
DATA RECORDS	.	.	.	.	.	.	.	.	.	.	.	.	.	.	->
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
	U	191Y	N	89	2	24	5	5	S	4	1.9	167	16	1.0	->
	U	191Y	N	90	2	24	5	5	S	4	1.9	167	16	1.0	->

C79	...	*C94*	*R72*	...	*R94*	*M72*	...	*M94*	*RUT86*	...	*RUT94*	*P79*	...	*P94*	*F79*	...	*F94*	->
00	...	05		...	49		...	62	0.12	...	0.20	00	...	00	5.0	...	4.0	->
00	...	15		...	57		...	45	0.29	...	0.30	00	...	00	4.5	...	4.0	->
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	->
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
05	...	12		...	112		...			...	0.20	00	...	00	5.0	...	4.0	->
03	...	00		...	123		...			...	0.15	00	...	00	5.0	...	4.0	->

MC_79	...	*MC_94*	*SN*	*PROJECT*	*YR*	*L1*	*T1*	*L2*	*T2*	...	*L6*	*T6*	*T*	*RATE*	*FY*
4	...	101	3.00	IIG 8- 1- 6	77	SM	12.0	AC	4.5	...			F	16.7	
1412	...	3202	3.00	IIG 8- 1- 6	77	SM	12.0	AC	4.5	...			F	30.3	
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
0	...	0	2.96	S207- - 12	61	SM	18.0	AB	3.0	...			F	28.8	
0	...	0	2.96	S207- - 12	61	SM	18.0	AB	3.0	...			F	77.0	

Note: ->: Continue.

Field Name Designations:

A: Functional Classification, I-Interstate, S=State Highway, U=US Highway;

RTNO, D, MP, L: Route Number, Direction, Milepost and Number of Lanes, respectively;

PW, LS, RS: Pavement Width, Left Shoulder Width, and Right Shoulder Width, respectively;

DIST, AREA, REG: District, Engineering Area, and Regional Factor, respectively;

ADT, ADL, GF: Average Daily Traffic, Average Daily Load (ESAL's), and Growth Factor, respectively;

Ci, Ri, Mi, RUTi, Pi, Fi, M_Ci: Cracking, Roughness, Mu Meter Number, Rutting, Patching, Flushing and Maintenance Cost for year i, respectively;

SN, Project, Yr: AASHTO Structural Number, Project Number, and Year of Last Project Applied, respectively;

Li, Ti, Rate, Fy: Type and Thickness of Layer i of Last Project, Priority Number and Fiscal Year respectively.

Figure 1.2. Data Format of the Pavement Management database for ADOT (after Wang et al., 1993).

Roughness

Roughness was the first pavement condition measurement used by ADOT; its collection dates to 1972. Roughness is recorded using 1 mile long intervals and the roughness value is assigned to the starting milepost.

Roughness is defined as "irregularities in the pavement surface that adversely affect ride quality, safety and maintenance costs" (FHWA, 1987). Roughness is a very important condition indicator because it correlates well with pavement serviceability, representing a good indicator of pavement riding confront. The concept of using ride quality or serviceability to evaluate pavement performance was one of the major accomplishments of the AASHO Road Test (AASHTO, 1990).

The concept is based on the principle that the prime function of the pavement is to serve the traveling public. The Present Serviceability Index (PSI) is an index that estimates the subjective reaction of road users to pavement quality. PSI uses a 0 to 5 scale, with 5 being excellent and 0 impassable. Serviceability or level of service can be determined as the average rating of a panel that rides the section and records the subjective opinions of the raters. The average opinion is known as the Present Serviceability Rating (PSR).

However, serviceability is normally estimated by correlation between the serviceability index and the calibrated output of the roughness device. In this case the resulting index is called the Serviceability Index (SI) (FHWA, 1985). ADOT has established an equation to compute the SI (ADOT, 1989):

$$SI = 0.3488 + 4.6836 * 0.9970^{(R-4.255)/0.54} \quad .....(1.1)$$

Where: R = calibrated Maysmeter number.

ADOT measured roughness at the network level using a Maysmeter until 1995. This is a response type roughness measurement device that is mounted on a full size car equipped with coil springs, firm shock absorbers and front and rear anti-roll bars. The Maysmeter measures the accumulated vertical movements, between the vehicle axle and frame, in terms of counts (tenth of an inch) and distance traveled by the vehicle. The counts are converted to inches per mile by dividing the total roughness counts by the distance traveled. A Maysmeter value is obtained by adjusting the measured inches per mile to account for vehicle calibration factors. A GMR-type profilometer (K.J. Law 690 DNC) was initially used for selected research sections and acceptance control of recently constructed projects. Starting 1996 the network level roughness data are collected with the GMR profilometer. The results of the GMR profilometer are reduced to Mays-meter values using the model defined in the National Cooperative Highway Research Project

(NCHRP) Report 228, "Calibration of Response-Type Road Roughness Measuring Systems" (Gillespie et al., 1980).

Pavement Surface Distress

The physical distress of a roadway section provides a measure of the deterioration caused by traffic, environment, and age. The types of distress and methods for evaluation vary widely from one agency to another (AASHTO, 1990). ADOT currently collects information regarding cracking, rutting, flushing and patching.

Cracking is measured as a percentage of a 1000 square foot area at each milepost. Rutting is measured as the mean depth of a rut in the wheel path, measured with a four-foot straight line edge. The K.J. Law Profilometer has been used for rut depth determination in selected research projects and special studies starting in 1991. Flushing, resulting from an excess of asphalt on the pavement surface, is measured using a severity rating index from 1 to 5 (1 indicates severe bleeding and 5 indicates no flushing). Patching is defined as any repaired or replaced area or surface treatment placed by maintenance forces. It is recorded as a percentage of the area of the pavement. The collection of cracking, flushing and patching data was started in 1979 and the rutting data dates to 1986. The defects are evaluated visually over a section 1,000 square foot over the travel lane starting at each milepost and are assigned to this milepost.

Skid Resistance

The principal pavement related factor that affects safety is the friction between the road surface and the tire (AASHTO, 1990). The friction or skid resistance can be measured objectively using automatic equipment. ADOT collects friction data, depending on the availability of manpower after condition and roughness surveys are accommodated, using a Mu-meter (Wang et. al., 1993). This device is a trailer that measures skid in a "yaw mode," where the wheels are turned at a 7.5 ° angle to the direction of travel. The Mu-meter measures the force required to maintain the yaw angle. The test is conducted over wet pavement by spraying a film of water with a calibrated tank-truck. Tests are generally made for 500 feet, beginning at the mile post.

1.2.3. Analysis

ADOT's PMS was oriented mainly to a program level analysis of the state's pavements using a Network Optimization System (NOS). This NOS determines the optimum work program and budget for ADOT's pavement network applying operations research techniques and, specifically, stochastic theory and optimization method. The output of the NOS is an *Optimum Budgetary Policy* that details the numbers of miles of pavements, within each road functional category, traffic level, and design region, that should receive each rehabilitation action, and the budget required to execute these actions. It also provides the projected percentage of pavements in each pavement condition group (Wang, et al., 1993).

Although NOS provides a very effective tool in generating annual budget needs, it does not provide information for selecting specific rehabilitation projects, due to the section aggregation method used in the NOS model. The Markovian prediction models used in NOS are based on the road categories rather than specific sites or projects. Therefore, they cannot be used as the tool for project selection.

The PMS does not automatically recommend specific rehabilitation actions for specific roadway sections. The project selection is conducted manually. The process starts with the preparation of a preliminary list of candidate projects obtained using an equation that computes a "Rate" number based on roughness, cracking, rutting, and maintenance costs. This preliminary list is reduced though meetings with the district engineers and windshield field evaluations of candidate sections. Based on the result of all additional considerations, the final list of projects to be included in the program is prepared.

Network Optimization System

The NOS provides a systematic, consistent, and theoretically sound method for determining the most cost-effective action for the different roadways in the state to achieve and maintain desirable maintenance standards over a specified planning period. Originally the NOS was designed to run in the mainframe computer environment. The overall NOS methodology contains the following components (Kulkarni et al., 1980):

- Selection of functional criteria and performance variables.
- Selection of influence variables for each performance variable.
- Selection of roadway categories and condition states.
- Specification of rehabilitation actions and policies.
- Pavement performance prediction.
- Optimization model for budget allocation.

The functional criteria describe the broad areas of concern that are relevant to determine the acceptability of pavement performance. The original study determined that safety, serviceability and physical distress were the most appropriate for use in the NOS. The performance variables, or measures of the degree to which the performance meets these functional criteria, are skid number, ride index, and amount of cracking respectively.

The study also determined that the main influence variables that affect the performance variables considered were: functional class, Average Daily Traffic (ADT), regional factor, index to first crack, present roughness, present amount of cracking, and change in amount of cracking during the previous year. Three of these factors were used to define road categories, functional class (Interstate and non-Interstate), traffic level (low, medium and high), and region within the state (mountains, transition and desert), resulting in 18 road categories in the original NOS. However, since low traffic levels do not exist for the Interstate roads, only 15 road categories were used. The other four variables, index to first crack, present roughness, present amount of cracking, and change in amount of cracking during the previous year, were used to define 120 pavement condition states (Kulkarni et al., 1980).

Originally the NOS considered 17 maintenance and rehabilitation options. However, not all actions are applied to all road categories and conditions, because some actions are not appropriate for some specific road conditions or categories.

The NOS also requires performance standards that define levels of pavement condition that are considered acceptable for the traveling public. A minimum portion of roads is required to be in a state of low roughness and low cracking and a maximum percentage of roads is permitted to be in a state of high roughness and/or high cracking (Kulkarni et al., 1980).

Pavement performance was modeled as a Markov process, in which the probability of transition from one condition state to another, in one year under one M&R action, gives the portion of roads in one state i that move to a state j in one year under this action.

The optimization model was formulated as a linear programming model. Its objective is to determine the minimum cost preservation program which maintains the network to the specific condition standards for each period (i.e., one year) of the analysis horizon. The model can be used to determine the optimum long-term rehabilitation policy if a stationary state is assumed, or to select a short term rehabilitation policy that would cause the system to transition from the current condition state to the desired steady-state. The flow chart of NOS analysis is shown in Figure 1.3.

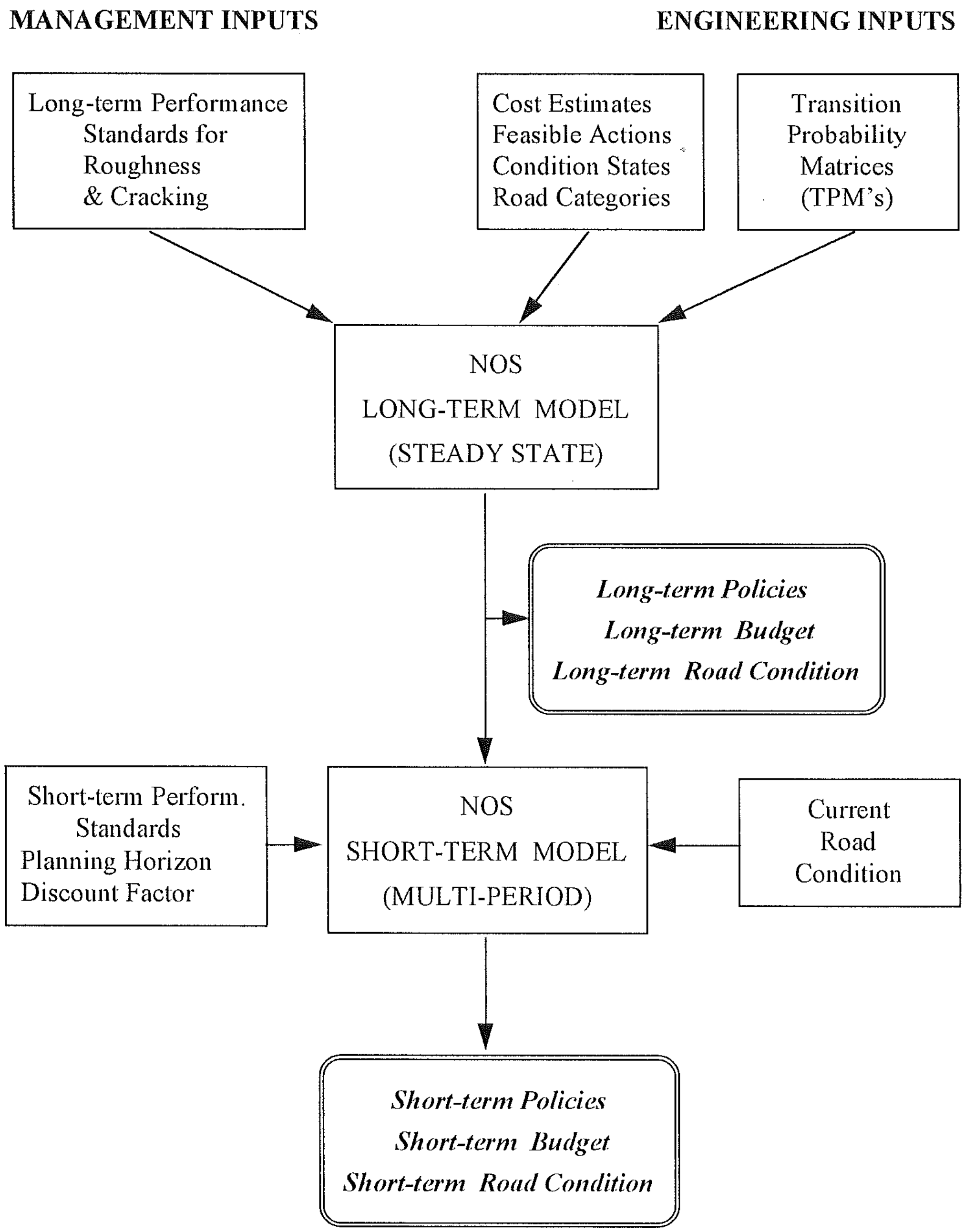


Figure 1.3. Flow Chart of the NOS Analysis (Wang et al. 1993).

Recent Revisions to the NOS

The NOS has been a very effective planning tool for the development for pavement preservation programs. The output of the NOS enables ADOT management to determine (Wang et al. 1993):

- What portion of the pavement in each road category will be expected to be in various condition states at the beginning of each time period,
- What is the most cost-effective rehabilitation policy for each roadway category in the network at each time period, and
- What are the expected annual cost of pavement rehabilitation and routine maintenance.

However, a comprehensive review of original mainframe-based version indicated that a few simplifications and improvements could be accomplished in order to make the model more tractable and effective. These improvements were implemented in new PC-based 32-bit linear optimizer (Wang, 1992).

Cracking change was removed from the system because the study showed that it was not significant in predicting the acceleration of pavement structural deterioration. An extensive evaluation of the cracking data showed that (Wang et al. 1993):

- The average cracking change over time was approximately lineal, contradicting the traditional assumption that the rate of deterioration accelerates after the initial cracking occurrence.
- The annual change in cracking rarely exceeded 5%, that is the error of the visual examination process used by ADOT.
- A large change in cracking, more than 5%, in one year does not indicate that it will be a high increase in cracking the next year.

The research effort also revealed that the effective number of preservation actions is 6 rather than the 17 used in the original system. Based on the effectiveness of the action in the year of application, there are three categories of actions:

- Routine maintenance (index to first crack = 0),
- Light treatments (index to first crack = 2), and
- Heavy treatments (index to first crack = 3 to 5).

All preservation treatments with the same index to first crack have the same transition probabilities under routine maintenance. Since all heavy treatments have approximately the same initial condition, a high probability of bringing the pavement to the best condition, all heavy treatments with the same index to first crack will have the same behavior. This restricts the NOS to select the least cost action for each heavy treatment group. Seal coats and asphalt concrete surface courses also have the same index to first crack. Although they are then grouped in the same category, they have different behavior because they have different initial conditions. Seal coats do not improve roughness (Wang et al. 1993).

Additionally, the levels of classifications for existing roughness and cracking needed to be adjusted to reflect current pavement conditions and standards. New transition probability matrices were generated based on past pavement performance database, and accessibility rules were established to ensure that pavements can only stay in their current condition or deteriorate to a worse condition.

The study included an extensive sensitivity analysis for the new NOS. It was revealed that the budget needs based on steady state runs are not necessary less than those based on multi-period runs. Therefore, the NOS multi-period runs should be used for the actual pavement preservation program, because the ideal situation represented by the steady state runs is not feasible (Wang et al. 1993).

The study recognized that although NOS provides a very effective tool for generating annual budget needs, it provides no capabilities for selecting rehabilitation projects, due to the section aggregation method used in the NOS model. Other approaches were needed to establish a comprehensive methodology for selecting rehabilitation projects. A computerized knowledge-based expert system was suggested, because it would allow a more detailed preliminary estimation of rehabilitation needs such that costs could be better ascertained (Wang et al. 1993).

Another concern regarding the system was that the total annual budget for the entire network is not imposed as a constraint in the current NOS. The answers from the existing NOS methodology are not applicable to the overall optimal rehabilitation for all 15 road categories if total annual budget constrains are imposed. In this case, a Global Network Optimization System is required. This can be achieved by maximizing a utility function representing network benefits, rather than minimizing agency cost. Benefit maximization is widely used in financial institutions in order to increase profit. Although the use of benefits as an objective for a public agency was recognized as a desirable feature, based on experience in ADOT, their use was considered impractical (Wang et al. 1993).

Finally, it was suggested that the integration of project selection procedures and global optimization in a single package within the framework of modular design and object-oriented programming could provide a comprehensive computer system for that would be able to conduct both financial planning and programming (Wang et al. 1994).

1.2.4. Existing Project Selection Procedure

The PMS does not automatically recommend specific rehabilitation actions for specific roadway sections. The NOS determines the optimum work program and budget for ADOT's pavement network, that details the numbers of miles of pavements, within each road functional category, traffic level and design region, that would receive each rehabilitation action, and the budget required to execute these actions. The allocation of the miles corresponding to each rehabilitation strategy, road functional category, traffic level and design region, to specific projects to be included in the Pavement Preservation Program is conducted manually.

The program is reviewed annually and the project list is adjusted as needed. The process identifies projects to be constructed starting five years after the collection of the available data, because it takes approximately 4 years from the time a project is selected to include in the plan and start of construction. For example, the projects constructed in fiscal year 1989 were selected in 1985 and using condition data collected in 1984.

The current selection process is summarized in Figure 1.4. It starts with the preparation of a preliminary list of candidate projects obtained using an equation that computes a "Rate" number based on the information about roughness, cracking, rutting, and maintenance costs stored in the PMS database. The formula used to compute the rate number is:

$$\text{Rate} = \text{Cr} + Rg/5 + 2*Rut + 0.0015*MC \qquad .....(1.2)$$

where: Cr = Cracking (%),

Rg = Roughness (Maysmeter units),

Rut = Rutting (inches), and

MC = Average maintenance cost for last 3 years (dollars per year).

The projects preliminary list is sent to each maintenance district along with a list of the projects that the District Engineers (DE) had requested the previous year which were not funded. The DE's consult with maintenance supervisors and discuss possible projects with the pavement management engineer to prepare a prioritized list of projects. This list may include projects which were not on the preliminary list. Generally, the DE's will add pavement sections that are maintenance concerns. Projects that have very high priorities according with the Districts' list, but were not included in the preliminary list, are generally added to prepare a "reviewed" preliminary list.

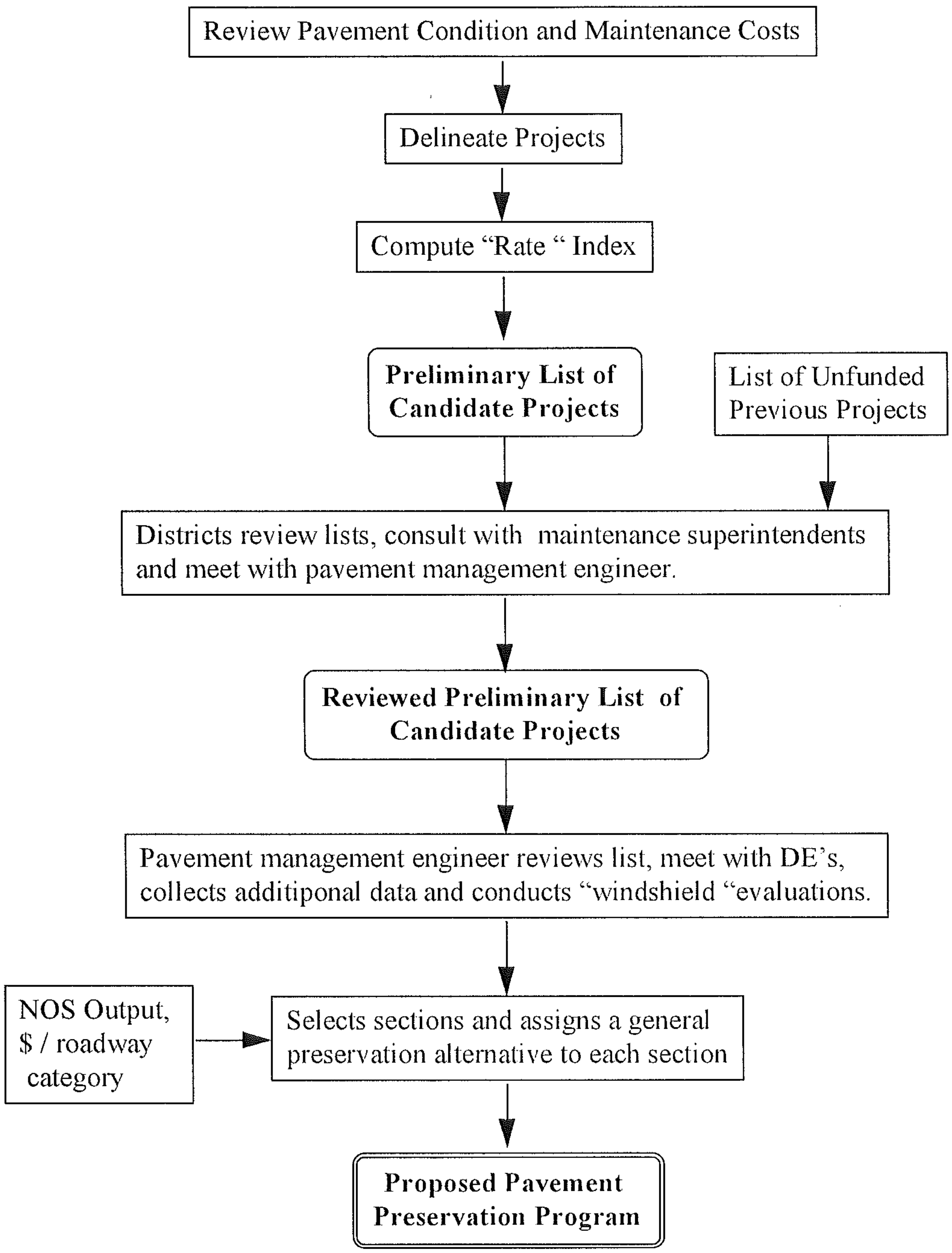


Figure 1.4. Current ADOT Project Selection Process.

The priorities of the reviewed preliminary list are further reviewed through meetings with the DE's and windshield field evaluations (*driving the section*) of particular candidate sections. At this stage, additional *format free* information is collected for a small portion of the network. This information is stored in hard copy form. Based in the result of all additional considerations, the final list of projects is prepared. For each rehabilitation strategy, road functional category, traffic level and design region the projects with the highest priority are selected until all the miles recommended by the NOS are allocated.

At this stage a "general" rehabilitation strategy (e.g., thin overlay) is assigned to each project. The most appropriate specific treatment design for each project is selected using life cycle cost analysis at the project design phase.

1.3. ISTEA REGULATIONS

The overall objective of the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 was the improved performance of statewide and metropolitan transportation systems through preservation, operational and capacity enhancements. This act required: "State development, establishment and implementation of systems for managing highway pavements of Federal-aid highways (PMS), bridges on and off Federal-aid highways (BMS), highway safety (SMS), traffic congestion (CMS), public transportation facilities and equipment (PTMS), and intermodal transportation facilities and systems (IMS)." The needs identified for these systems had to be considered in developing metropolitan and statewide transportation plans and improvement programs. Therefore, the role of the management systems was both, the development of information and strategies to improve the performance of the existing and future facilities and to provide input to the planning process for consideration at the system level.

The general requirements for the State PMS's were:

1. Each State shall have a PMS for Federal-aid highways.
2. The States are responsible to assure that all Federal-aid highways in the State, except those federally owned are covered by a PMS.
3. The PMS's should be based in the AASHTO Guidelines for Pavement Management Systems.
4. Pavements shall be designed to accommodate current and predicted traffic needs in a safe, durable, and cost-effective manner.

The regulation also established that the PMS for the National Highway System (NHS) shall, as a minimum included the following components:

1. Data Collection and Management.
   i. Inventory of physical pavement features, including the number of lanes, length, width, surface type, functional classification, and shoulder information.
   ii. History of project dates and types of construction, rehabilitation, and preventive maintenance.
   iii. Condition surveys that include ride, distress, and surface friction.
   iv. Traffic information including volumes, classification, and load data.
   v. A database that links all files related to PMS. The database shall be the source of pavement related information reported to FHWA for the Highway Performance Monitoring System (HPMS) in accordance with the HPMS Field Manual (FHWA, 1993).
2. Analysis.
   i. A pavement condition analysis that includes ride, distress, and surface friction.
   ii. A pavement performance analysis that includes an estimate of present and predicted performance of specific pavement types and an estimate of the remaining service life of all pavements on the network.
   iii. An investment analysis that includes:
      a. A network-level analysis that estimates total costs for present and projected conditions across the network.
      b. A project-level analysis that determines investment strategies including a prioritized list of recommended candidate projects with a recommended preservation treatment that span single-year and multi-year period using life-cycle cost analysis.
      c. Appropriate horizons, as determined by the State, for these investment analyses.
   iv. For appropriate sections, an engineering analysis that includes the evaluation of design construction, rehabilitation, materials, mix designs, and preventive maintenance as they relate to the performance of pavements.
3. Update.

   The PMS shall be evaluated annually based on the agency's current policies, engineering criteria, practices, and experience, and updated as necessary.

The National Highway System Act, November 1995, made some of the ISTEA requirements optional. However, ADOT engineers and authorities believe in the cots-effectiveness of sound roadway management practices. Therefore, ADOT has maintained its commitment towards the implementation of the proposed management systems.

1.4. ENHANCEMENTS NEEDED TO MEET NHS PMS REQUIREMENTS

ADOT PMS, in general, met the requirements for NHS PMS's. However, the system, and mainly the analysis module, needed some enhancements to fully comply with the federal requirements.

The most important possible enhancements to the system were:

- Review the performance curves for the different pavement types and develop a methodology to estimate Remaining Service Life (RSL) of all pavement on the network.
- Develop a computerized project selection level analysis procedure for project recommendation and prioritization. This procedure must recommend a list of candidate projects to include in the pavement preservation program.

These two points appeared necessary because the Interim Final Rule explicitly required that the PMS must include the analysis of RSL for the NHS and must produce a recommended list of projects. The automatic project recommendation procedure, using RSL and other considerations, must provide a prioritized list of candidate projects to be included in the rehabilitation program. At this stage a "general" rehabilitation treatment (e.g., thin overlay) is assigned to each project. The sections to be included in the preservation program can be selected based on this list. Then, the most appropriate specific treatment design for each project can be selected using life cycle cost analysis at the design level.

A research project consisting of a long-term follow-up of the performance of selected sections may be necessary. However, this point could be covered by several on going activities, such as the SHRP's Long Term Pavement Performance (LTPP) and SPS sections as well as other long term studies of pavement performance conducted by ADOT.

1.5. OBJECTIVE OF THE RESEARCH

The objective of the research was to developed and implement an automatic Project Recommendation Procedure (PRP) for ADOT's PMS. This procedure provides recommendations for projects that should be considered for the pavement preservation program. The PRP had to use the most appropriate technology available and consider the RSL of all pavement sections.

The output of the project recommendation procedure is a list of projects, sorted by priority, for each rehabilitation strategy, road functional category, traffic level and design region. Such a procedure makes ADOT's PMS more project selection oriented. The final program selection would still be conducted manually.

The main components that were identified as necessary for the project recommendation process were:

1. Pavement section delineation.
2. Pavement performance prediction.
3. Remaining service life analysis.
4. Preliminary screening of the network and selection of candidate sections.
5. Project recommendation (including computation of priority index and treatment recommendation).

The development of a user-friendly shell, that integrates database management, project recommendation, and report generation, was defined as a secondary goal of the research effort.

1.6. RESEARCH APPROACH

The research effort started with a comprehensive literature review that included the program development process, section delineation techniques, pavement performance prediction, remaining service life analysis, project selection and prioritization methods, and pavement management artificial intelligence applications.

The first step in the development process was the definition and implementation of a section delineation procedure that process the milepost data into uniform roadway sections or management units. The selected delineation procedure starts by dividing sections in accordance with the last construction project applied to each milepost. The subsections longer than a minimum length are further divided, using the cumulative difference method, if they have points of significant condition change. The average roadway characteristics and pavement condition indicators for each uniform section are then computed.

Linear regression models were selected to determine project selection level pavement performance equations. A computer program that automatically determines the best-fit roughness and cracking linear equations for each uniform section was implemented.

"Trigger points," or threshold values at which the pavements were considered in need of a rehabilitation treatment, were defined for several condition indicators based on the opinion of a selected group of ADOT experts. The average opinion of the experts was determined using a survey questionnaire.

A program that computes the remaining service life of each uniform section was developed. It computes the remaining service life as the minimum of the remaining lives due to roughness and cracking determined using the performance prediction equation and trigger point for each indicator independently. A maximum service life was established for each roadway functional classification and preservation treatment, using the 95 percentile of the distribution of service lives of all treatment in the projects database.

An artificial neural network simulator was used for the preliminary screening of candidate projects. The simulator was trained with the pavement condition and characteristics at the time of programming, and the sections selected for the pavement preservation program. The most appropriate network architecture was determined using a designed fractional factorial experiment. The experiment identified the factors that significantly affect the network performance. The final network architecture was defined by analyzing the significant factors to determine their most appropriate settings. The trained artificial neural network selects candidate roadway sections using the criteria learned and the current condition of all pavement sections in the network. In addition to the sections recommended by the artificial neural network, all sections with remaining service life less than 3 years, excessive maintenance cost or high "rate numbers" are included in a preliminary list of candidate sections.

The preliminary list of candidate sections is further analyzed by an "expert" project recommendation computer program that recommends a preservation treatment and assigns a priority rating to each section in the list. The knowledge required to develop the recommendation procedure was obtained from a selected group of ADOT experts using a survey. The survey questionnaire was designed using the fractional factorial methodology. The survey responses were used to determine a priority equation and a treatment recommendation formula. However, treatment recommendation formula was only used as a reference to define a treatment assignation matrix, because important factors, such as the structural number, were not significant statistically in the model equation. The recommendation program sorts the projects by priority and assigns funding to the highest priority sections within each roadway group until the budget recommendation provided by the network optimization process is reached.

The comprehensive project recommendation procedure was implemented in a user friendly, menu-driven, Windows-based, modular computer program. The program provides access to the databases, conducts the entire project recommendation process (or parts of it), displays the results, and prepares and print reports.

2. LITERATURE REVIEW

2.1. PROGRAM DEVELOPMENT

A highway preservation program is composed of projects to be rehabilitated and specifies when and how to rehabilitate each project. The National Cooperative Highway Research Program Synthesis of Highway Practice Number 48 defines programming as "the matching of available projects with available funds to accomplish the goals of a given period." It is important to distinguish between planning and programming. Planning deals with long-range goals and objectives, idealistic, unconstrained and policy-oriented plans. On the other hand, programming is associated with short-term, fiscal constrains, priorities, and project-oriented programs (NCHRP, 1978). ADOT's current network optimization system is a planning tool; the project recommendation procedure developed in this research is a programming tools.

2.1.1. The Programming Process

The NCHRP Synthesis Number 48 (NCHRP, 1978) reported that there was no generally recognized format for the basic programming process. However, it identified a 15-step process, that can be considered appropriate for describing the programming process:

1. Project initiation - the projects are obtained from two major sources: technical (planning studies, condition deficiencies) and non technical (request and observations).
2. Initial listing - including all candidate projects identified that can be originated at several levels within the agency, i.e., headquarters, districts, counties.
3. Preliminary analysis - projects are analyzed on the basis of existing or easily obtainable information.
4. Combined listing (first draft) - any separate lists are combined in a single list.
5. Advanced analysis and prioritizing - once all projects are identified and listed, they need to be thoroughly analyzed and prioritized. This step has three major components: technical prioritizing, non technical prioritizing, and feedback from planning and development.
6. Combined listing (second draft).
7. Financial analysis - the analysis should address two fundamental issues: how much is available and how much is committed.

8. Preliminary program - the programmer defines the first program by matching the list of prioritized project with the available funds.
9. Executive session - the top staff of the department head holds a meeting to select a program for publication.
10. Short-range program (first draft) - a short-term program is published.
11. Executive and legislative review - the first draft of the program is reviewed with administrative staff and legislative leadership.
12. Short-range program (final draft) - after review, and modification if necessary, the program becomes official.
13. Scheduling - a detailed schedule for all phases, from planning to construction, is developed.
14. Monitoring - the actual progress is compared with the scheduled progress.
15. Modifying - the program is amended as necessary.

All fifteen points are related, to some degree, with the objective of this research. However, the project recommendation procedure will mainly concentrate in steps 1 to 8. In particular the fifth step, analysis and prioritization of projects, warrants further consideration.

Both, technical and non technical elements are combined in the decision-making process. The implication of technical prioritizing is that a person or staff having technical training could evaluate all important factors and compute a "magic number" for ranking projects. This would completely remove politics, with its negative connotation, from influencing the spending of available transportation funds. The technical factors that were considered in prioritizing, at the time the synthesis was prepared, were:

1. **Sufficiency ratings**, are the earliest and simplest ratings, and are based on the structure, serviceability, and safety of the section of roadway.
2. **Priority ratings**, represent the next generation of priority rating, and incorporate other factors, such as traffic volumes, accidents, as well as social, economic and environment factors.
3. **Option-evaluation techniques**, comprehensive or sketch planning. They represent the third generation of prioritizing techniques that attempt to be all-encompassing.

On the other hand, non technical prioritization includes items that cannot be measured or are difficult to quantify objectively. The weighing of their importance is done in the minds of the decision makers. Some of the elements considered are political commitments, legislative mandate, emergency, special emphasis, commitments to other agencies, system continuity-connectivity (missing links), and position in pipeline (project readiness).

An important concept identified in the synthesis is the need for a definite programming process that is fully understood by everyone who may affect it. Understanding the process helps prevent delays in the project development and encourages stabilized budget funds (NCHRP, 1978).

2.1.2. Single and Multi Year Programs

Kher (1991) suggested that the program development process can be divided in two levels:

1. Single year program development, and
2. Multi year program development.

Single Year Program Development

The single year program development attempts to develop a single year program which optimizes expenditures for the current year. It involves technical, program development and management aspects, focusing in the short term end of the program development process.

The first step in the single year program development process is to determine what pavement sections should be considered for preservation. The second step is to determine which factors will be used to identify a project that is in need for preservation, such as the project overall condition, individual distresses, rate of deterioration, and maintenance costs. Next, "trigger points" or threshold values must be defined. These trigger points indicate the level of a particular performance factor at which the agency considers a pavement a candidate for rehabilitation or preventive maintenance. Trigger points may vary according to the importance, use, or classification of a facility, as indicated in Figure 2.1.

Candidate pavements are triggered for treatment either now or at some point in the future and the network is categorized into "current needs" and "future needs." Once the current needs are identified the appropriate treatment and cost for each section must be determined. The projects are prioritized according to an appropriate criterion that must be in concordance with the agency overall goals and objectives. The single year program can be prepared based on the prioritized list.

Multi Year Program Development

The single year program is a subset of the multi year program. The concepts of single year program development can be generalized for multi year program development if the system has the capability of predicting the condition of the pavements in the future. The approaches for multi year program development range from simple ranking to the use of mathematical optimization techniques (Haas, 1991).

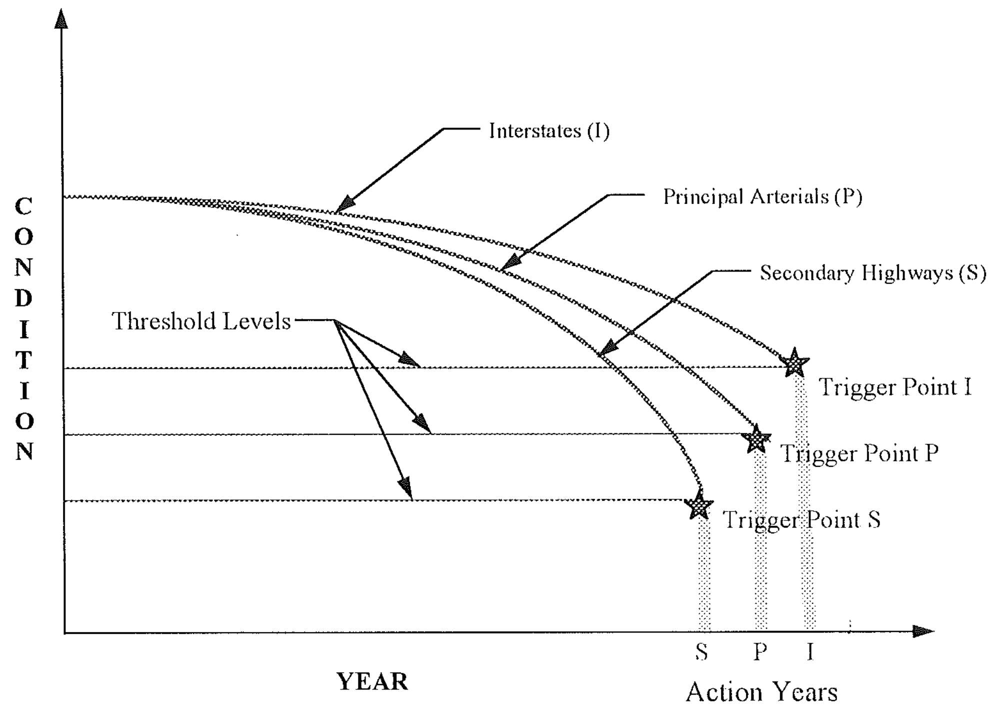


Figure 2.1. Example of Trigger Points for Different Road Characteristics (Kher, 1991).

The multi year program development should seek to achieve the best combination, over a specified program period, of:

1. Where are treatments needed, or which sections in the network are to be maintained, rehabilitated or reconstructed.
2. What maintenance, rehabilitation or reconstruction treatments, from the alternatives available, are to be applied to those sections.
3. When should the treatment be carried out.

The main advantages of multi year programming with respect to single year programming are (Haas, 1991):

1. Timing of M&R can be included in the process.
2. Allows an optimum combination of projects-alternatives-timing, for any specified budget level.
3. Can be used for strategic purposes, for example setting targets for future, such as maximum number of miles under a specific serviceability level.

4. Allows assessment of the effects of different funding levels.
5. Allows identification of a future year for rehabilitation of a specific section, in response to sequential queries.

The only disadvantages of multi-year programming are that it requires reliable performance prediction models and that it is more complex than single year prioritization. Although reliable performance prediction models enhance the capabilities of the PMS, some agencies can consider the need of such models a disadvantage because they are difficult to obtain. Also the development of performance prediction models requires reliable pavement performance records that are not always available.

2.2. DETERMINATION OF HOMOGENOUS SECTIONS

Given the milepost-based structure of ADOT pavement database, the first step in the project selection process is to delineate roadway project units or sections with approximately homogenous physical characteristics and condition. For any preservation treatment, it is more cost-effective to rehabilitate projects that are homogenous in condition than to rehabilitate projects that do not have uniform conditions. For this reason, the entire network should be partitioned into section in such a way that the pavements in each section are uniform in terms of pavement type, condition and functional classification (Kuo et al., 1992). Each uniform or homogeneous unit should be considered as a candidate project.

The uniform units or sections should have not only the same pavement type, construction history (including M&R), cross section, subgrade and traffic, but also relatively low variability of the response variables, such as deflection, SI, skid number, pavement condition indices or individual distresses, along the section (AASHTO, 1993).

The measurement of a response variable changes from one location to another along a highway segment, with some points experiencing changes of significant magnitude. At these points of significant change, the overall response of the pavement on either side will be noticeable, as illustrated in Figure 2.2.

As shown in Figure 2.2, the variability when measuring pavement response can be traced to two major sources:

1. Between unit variability and
2. Within unit variability.

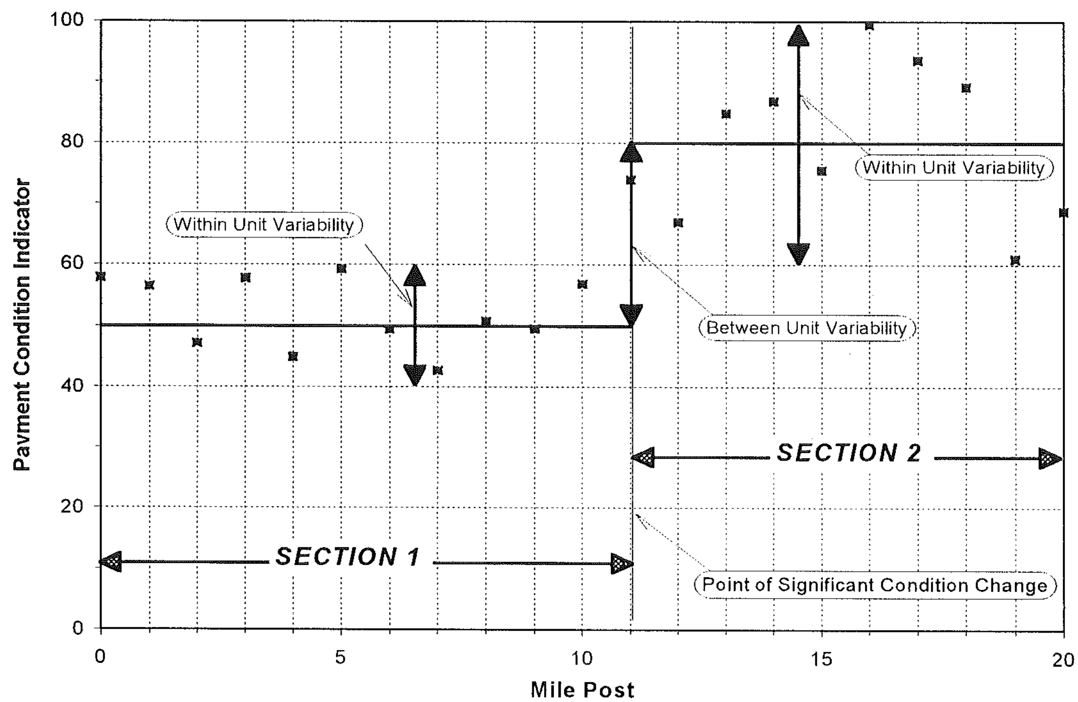


Figure 2.2. Example of Section Delineation Based on Significant Changes of Pavement Condition.

The first source of variation, or between unit variability, reflects the fact that several statistically homogenous sections may exist within a roadway segment. The determination of the boundaries between the different statistically homogenous sections is important in the project selection process because different rehabilitation strategies may be warranted. The within section unit variability is given by the inherent variability of the response variable within each section (AASHTO, 1993).

2.2.1. Unit Delineation Methods

The AASHTO Guide for Design of Pavement Structure (1993) distinguishes between two methods of unit delineation:

1. Idealized approach, and
2. Measured pavement response approach

In the idealized approach the engineer should isolate each unique factor influencing potential pavement performance, such as pavement type, construction and maintenance history, pavement cross section (layer materials and thickness), subgrade, traffic and pavement condition. Using historical data, sections that are characterized by a unique combination of all pavement performance factors can be determined. The validity of the final units is directly related with the accuracy of the historical data.

The measured pavement response approach relies upon the analysis of a measured pavement response variable (e.g., deflection) for unit delineation. A plot of the measured response variable as a function of the distance along the project can be used to delineate units using several methods. The simplest method is visual examination to subjectively determine relatively homogenous units. In addition, several analytical procedures are available.

Kuo et al. (1992) suggests the *Automatic Interaction Detection* method. This method uses analysis of variance techniques to find the number of uniform sections that are maximally significant. The rules such as minimum and maximum lengths of a uniform section, as well as, other practical and engineering rules are part of the procedure (Sonquist and Morgan, 1964).

AASHTO (1993) suggests the *Cumulative Difference* procedure. This procedure uses a variable, Z_x, defined as the difference between the area under the response curve at any distance and the total area developed by the overall project average response at the same distance. When this variable Z_x is plotted as a function of distance along the project, unit boundaries occur at the locations where the slopes (Z_x vs. x) change sign. This unit delineation methodology is used by MODULUS, a back-calculation program developed by the Texas Transportation Institute (Uzan et al., 1988), to identify sections with uniform pavement deflections. Although this method was formulated to delineate section based on pavement response measurements, it can be generalized to use other pavement condition indicators, such as roughness and cracking.

2.2.2. Unit Delineation by Cumulative Differences

The overall approach for this procedure is shown in Figure 2.3. In this example, the response variable is continuous and has three constant levels (r_1, r_2, and r_3) within three intervals along a project length. Therefore, there has to be three uniform sections having different response magnitudes as indicated in the first plot. The overall average response for the entire project is shown as a dashed line. In the second plot, the solid line shows the trend of the cumulative area (A_x) under the response variable r, and the dashed line represents the cumulative area ($\overline{A}_x$) caused by the overall average project response. The cumulative Z_x is determined as the difference of these areas:

$$Z_x = A_x - \overline{A}_x \qquad .....(2.1)$$

If the value of Z_x is plotted against distance, as shown in the third plot, the locations of the unit boundaries always coincide with the local maximums and minimums of Z_x (where the slope change signs). This fundamental concept is the basis used to analytically determine the boundary locations.

The basic concept can be extended to a discontinuous response variable r_i measured at every milepost using a numerical difference approach. In this case, Z_x can be computed as:

$$Z_x = \sum_{i=1}^{n} a_i - \frac{n}{n_t} \sum_{i=1}^{n_t} a_i \qquad .....(2.2)$$

where Z_x = cumulative difference at milepost x,

n = number of milepost measurements conducted from the beginning of the roadway section under study,

n_t = total number of milepost measurements conducted on the roadway section under study,

a_i = average of the response variable between mileposts i and I+1.

For the response that are measured in a specific location at the milepost (e.g. cracking), the average response, a_i, for the roadway mile can be computed as:

$$a_i = \frac{r_{i+1} + r_i}{2} \qquad .....(2.3)$$

where r_i = response measurement at milepost i,

r_{i+1} = response measurement at milepost I+1

2.2.3. Additional Considerations

After the units are determined, minimum and maximum length constraints, as well as other practical and engineering rules must be applied. The length of each unit to determine whether it has to be further divided or if two or more units should be combined for practical construction and administrative considerations and/or economic reasons. The mean response of each unit, and its influence on future performance of rehabilitated pavements, must be considered before combining the units.

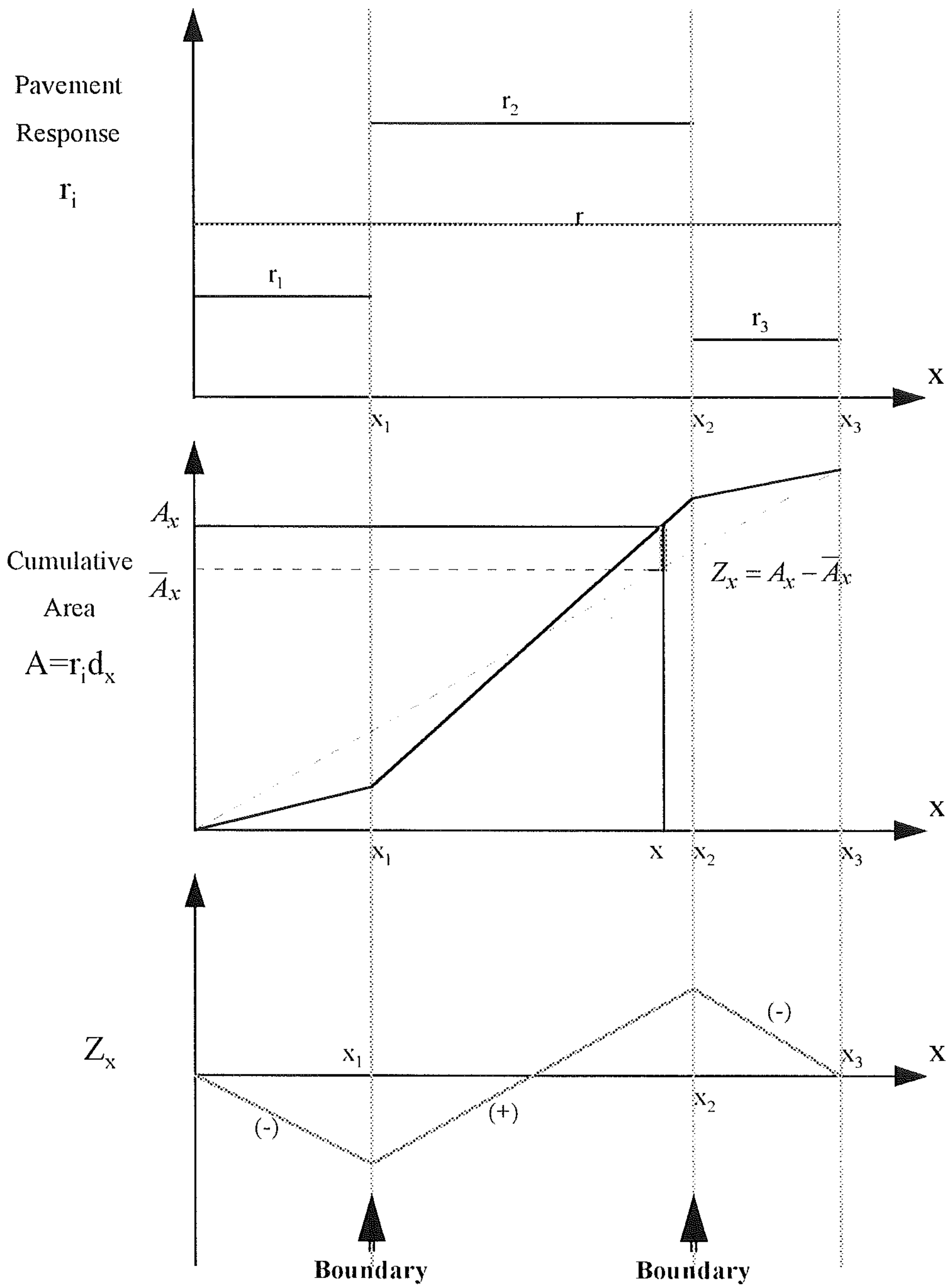


Figure 2.3. Concept of Cumulative Difference Approach for Unit Delineation (AASHTO, 1993).

The best solution appears to be a delineation procedure that combines the idealized approach, based on pavement characteristics and historical data, with the measured pavement performance approach to determine the final section delineation.

2.3. PAVEMENT PERFORMANCE PREDICTION

Prediction of the pavement condition or performance over time, although difficult, allows to use more sophisticated programming techniques. According to Lytton (1987) the capability of modeling pavement performance is a very important feature of both project and network level pavement management.

Important reasons to include prediction models in the pavement management process are (Mahoney, 1991):

1. To predict future condition of a roadway section.
2. To estimate the type and timing of M&R treatments.
3. To optimize the pavement condition for a complete highway network.
4. To use as a "feedback" loop to the pavement design process.
5. To use in pavement life-cycle cost analysis.

Darter (1980) defined the basic requirements for developing a condition prediction model:

1. An adequate database.
2. Inclusion of all significant variables affecting deterioration.
3. Careful selection of the functional form of the model to represent the physical relationships in a real word situation.
4. Criteria to assess the precision of the model.

There are several types of performance models that can be separated in two groups (Lytton, 1987):

1. **Deterministic** models and
2. **Probabilistic** models.

Deterministic models are generally regression models that can predict a single value of some parameter. On the other hand, probabilistic models predict a range of values of some parameter. The deterministic models have been used to predict (Lytton, 1987):

1. Primarily response (e.g., deflection, stress, strain).
2. Physical condition (e.g., distress, pavement condition index).
3. Functional parameters (e.g., PSI, safety).

4. Damage (e.g., load equivalents).

The probabilistic methods include survivor curves and transition processes, such as Markov and Semi-Markov models. While deterministic prediction models are used for both network and project level of PMS, probabilistic models are used only for the network level.

2.3.1. Methods to Develop Performance Models

Broadly speaking, pavement performance models can be developed using the following techniques (Mahoney, 1991):

1. Regression Analysis
2. Markov Transition Probabilities, or
3. Bayesian Methodology.

In the regression analysis, a dependent variable of observed or measured functional or structural deterioration is related to a one or more independent variables like subgrade strength, pavement layer thickness, traffic and environmental factors. If the least squares method is used, as is usually the case, the best relationship is the one that minimizes the differences between the values predicted using the regression equation and the actual data.

In the Markov process, the future state of the modeled element, a pavement section, is estimated solely from the current state of the element. The state of the element is defined by condition measures such as serviceability, pavement condition indices, cracking, skid number, etc. A transition probability matrix defines the probability that a pavement in an initial condition state will be in some future condition state. An example of one of the transition probability matrices currently used by ADOT for network level analysis is shown in Figure 2.4. According with this transition probability matrix, 72% of the pavement in state 2 will remain in state 2 next year, 15% will deteriorate to sate 3, 8% to state 5, and 5% will be in condition state 6 next year.

The Bayesian Methodology allows the analyst to combine both subjectively and objectively obtained data to develop a predictive equation. In fact, the approach can be used to produce regression equations exclusively developed from subjective information, as in the NCHRP Project 9-4 (Smith et al., 1979). In Bayesian regression analysis the regression parameters are assumed to be random variables with associated probability distributions.

To State	1	2	3	4	5	6	7	8	9
From State									
1	0.85	0.06	0.00	0.08	0.01	0.00	0.00	0.00	0.00
2	0.00	0.72	0.15	0.00	0.08	0.05	0.00	0.00	0.00
3	0.00	0.00	0.74	0.00	0.00	0.26	0.00	0.00	0.00
4	0.00	0.00	0.00	0.87	0.06	0.00	0.07	0.00	0.00
5	0.00	0.00	0.00	0.00	0.64	0.16	0.00	0.04	0.16
6	0.00	0.00	0.00	0.00	0.00	0.87	0.00	0.00	0.13
7	0.00	0.00	0.00	0.00	0.00	0.00	0.94	0.06	0.00
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.95	0.05
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Figure 2.4. Transition Probability Matrix under Routine Maintenance, High Traffic Road Category of Interstate at Desert Region (Wang et al., 1993).

2.3.2 Project vs. Network Level Prediction Models

The development and use of a pavement performance model depend on whether the models are used at the project or network level. The following aspects may be considered (Mahoney, 1991):

1. The availability of appropriate data to develop a pavement model is critical. This is not generally a problem at the project level because project specific information, such as annual condition rating, age, ESAL's, layer thickness, etc., is usually available. However, network level models often require information regarding several independent variables to account for the different characteristics of the sections in the network, and this information is sometimes difficult to obtain.
2. Project level models, in general, model and predict better than network level models.
3. The model form will also be influenced by the level of application. Usually, network level models will require transformations of the variables and can present multicolinerity problems because of correlation between the independent variables.

The considerations regarding information needs and availability and model complexity for project selection level performance predictions lie somewhere between the two extremes described above. A project selection level performance analysis requires models of complexity similar to those required for a project level but they have to be defined based on less detailed information. Therefore, the predictions at the project selection level are less precise than those obtained at the project-level.

2.4. REMAINING SERVICE LIFE

The Remaining Service Life (RSL) of a pavement section is a very important pavement performance indicator that should be taken into consideration in the project selection process. Baladi (1991) defines RSL of a pavement section as the estimated number of years, from any given date (usually from the last survey date), for a pavement section to accumulate distress points equal to a defined threshold value. The RSL of a newly designed and constructed or rehabilitated pavement section is equal to the design life. No negative remaining service life should be assigned to any pavement regardless of its condition. Any pavement section that falls bellow the threshold value has a zero remaining life. The RSL of a network is the average, weighted by length, of the remaining service lives of all pavement sections within the network.

The RSL concept is closely related with the "trigger point" or condition threshold value as defined in section 2.1.2. The definition of the threshold values depends on the criteria used to control long term network condition. Some of the criteria that can be used to establish the threshold values are (Kuo et al., 1992):

1. The value at which a M&R treatment is needed to remove unacceptable or poor condition.
2. The value at which routine maintenance is required to maintain pavement serviceability.
3. The value at which a M&R treatment will cost the agency and users less for rehabilitation than it would cost them if no rehabilitation is undertaken.

A pavement with a zero RSL does not mean that the pavement is "dead or gone." It simply means that one of the condition indicators used in its calculation has reached the threshold value defined for the RSL analysis (Kuo et al., 1992).

The RSL can be useful as an indicator of current and future pavement rehabilitation needs, to assess the impact of various budget levels on the health of the network, and to plan yearly balanced programs.

2.4.1. Justification

Any itemized, combined, or overall pavement condition index used to rate pavement condition has the limitation of not expressing the rate of deterioration. Two pavements having the same condition index may or may not require rehabilitation at the same time, depending of their rate of deterioration. Additionally, the combined or overall indices, computed on an averaging process, cannot be used to determine feasible rehabilitation alternatives, nor to asses the damage contributed by each distress mechanism.

On the other hand, the RSL of a pavement combines the severity and extent of the different distresses and the rate of deterioration. The RSL requires development of a performance model and establishment of a threshold value for each distress type. Based on these threshold values, the current distress level, and the deterioration model for each particular distress, the time for each distress to reach the threshold value can be computed. The shortest of these time periods is the remaining service life of the pavement section (Baladi, 1991). The RSL concept is illustrated in Figure 2.5.

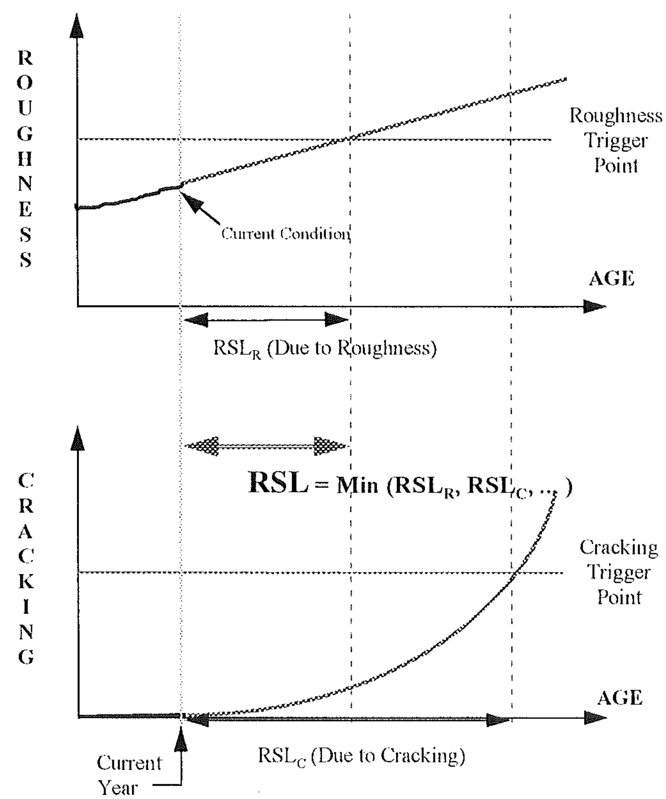


Figure 2.5. Remaining Service Life Concept.

2.4.2. Uses and Limitations

The RSL concept can be useful for several applications. Baladi (1991) highlights the following uses:

1. Estimate the remaining predicted service life of the various sections and the average remaining service life of the pavement network.
2. Calculate the percentage of the network in each RSL category or group (e.g. 0-5, 5-10, 5-15, 15-20, and < 20 years).
3. Detect any uneven distribution in the RSL of the pavement networks.
4. Estimate the health of the network in terms of lane-mile-year in acceptable condition.
5. Estimate the benefits of rehabilitation in terms of lane-mile-year gained.
6. Select an "optimum" rehabilitation alternative based on maximizing the RSL of the network or balancing the distribution of percentage of the network in each RSL group.
7. Generate multi-year rehabilitation programs.
8. Asses the impact of various budget levels on the health of the network (in terms of lane-mile-year gained or lost).
9. Enhance communications to legislators concerning network needs.
10. Control further conditions of the pavement network.
11. Calculate the percentage of the network users traveling on substandard roads (zero RSL).
12. Allow the planning of yearly balanced programs.

On the other hand, Baladi (1991) recognizes the following elements as limitation of the RSL concept:

1. Difficulty to establish the threshold values for each particular distress (definition of failure).
2. Difficulty to determine rate of deterioration.
3. Some distress type must be considered separately for specific reasons. For example, a blow up in a concrete pavement would require immediate action.

An additional limitation is that a pavement section with a zero RSL does not necessarily mean that the pavement is not longer in service. It simply means that one of the condition indicators used in its calculation has reached the threshold value defined for the RSL analysis. A pavement section with zero RSL may be in service for many years, if it is properly maintained.

2.5. TREATMENT SELECTION AND PROJECT PRIORITIZATION

Treatment selection for each candidate uniform roadway unit, and prioritization among the resulting projects, are two important steps in the programming process.

2.5.1. Treatment Selection Methods

The criteria used to select the most appropriate preservation treatment for a roadway segment should be based in at least one of the following factors:

1. Overall project condition (e.g., PSI or an overall surface distress index),
2. Individual distress,
3. Rate of deterioration, and
4. Maintenance cost.

The most appropriate criteria for a particular highway will depend on the agency requirements, availability of resources, and sophistication desired. According to Kher (1991) there are six methods of treatment selection:

1. **Policy and Experience**. A logic schema, that assigns the "most successful" M&R action to each pavement condition, road category and traffic, can be used.
2. **Decision Trees**. The selection is based on a series of criteria relative to individual distress type, condition indicators, and other factors, that reduce to the final recommendation. The decision trees are essentially similar to the first method. Figure 2.6 shows an example of a decision tree for selecting M&R treatment for flexible pavements.

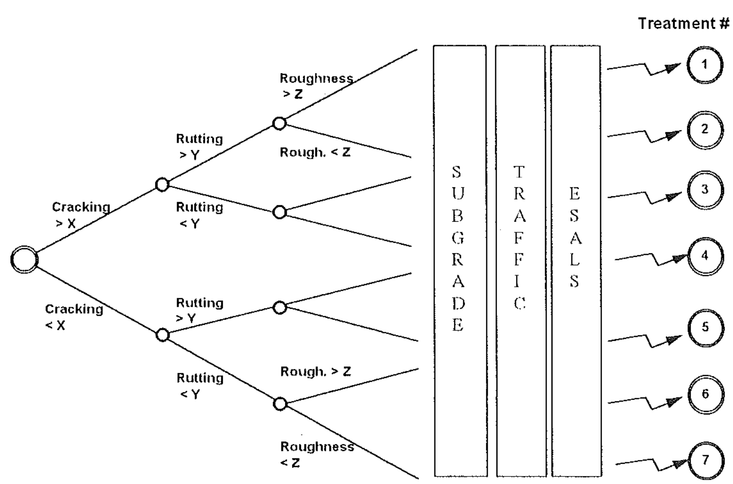


Figure 2.6. Example of Decision Tree for Selection of M&R Strategies (Kher, 1991).

3. **Design Methods (Mechanistic or Empirical)**. They can be used to determine appropriate treatments such as thickness of overlay to individual project. These procedures are more appropriate for a project-level pavement management.
4. **First Cost**. Sometimes, when the budget is very limited, the manager selects the treatment with least initial cost. However, this procedure it is not recommended because it does not consider future consequences, such as a short life of the treatment or the potential need for excessive maintenance.
5. **Least Present Cost**. It is based in a life-cycle cost economic analysis. Several possible treatments are considered for each section. All the cost incurred during a fixed period of time, or analysis period, are computed and discounted to the analysis year using a discount rate. These costs are summed to get the Total Present Cost (TPC) and the projects with the lowest TPC are preferred.
6. **Benefit/Cost and Cost-Effectiveness Analysis**. Both techniques involve economic analysis. The Benefit/Cost (B/C) ratio method uses the same analysis procedure as in the least present cost method to compute benefits and cost of each M&R strategy. The project with the highest B/C ratio should be selected. The Cost-Effectiveness analysis computes the benefit of an alternative in terms of effectiveness instead of monetary benefits. In some cases, effectiveness is defined as the area between the performance curve for a treatment and the "do nothing" curve as shown in Figure 2.7.

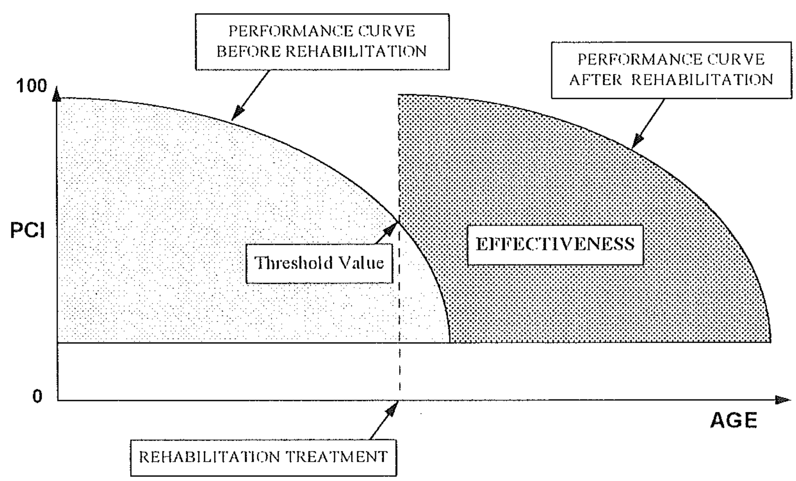


Figure 2.7. Effectiveness Concept.

Life-cycle Cost Analysis

Several effective project selection methods are based on a life-cycle cost economic analysis. Life-cycle costing can be defined as "an economic assessment of an item area, system or facility and competing design alternatives considering all significant cost of ownership over the economic life, expressed in terms of equivalent dollars" (Dell'Isola and Kirk, 1981).

The costs that a public agency may consider in the economic analysis include the following (Haas et al., 1994):

1. **Agency costs**: initial capital cost of construction, future capital cost of rehabilitation or reconstruction, maintenance costs, salvage or residual value at the end of the analysis period (which may be a negative cost), engineering and administration costs, etc.
2. **User costs**: travel time, vehicle operation, accidents, discomfort, and time delay and extra vehicle operation costs during resurfacing or major maintenance.
3. **Nonuser costs**: air pollution, noise pollution, neighborhood disruption.

Several economic analysis methods have been used to compare alternatives. Present-worth, equivalent uniform annual cost, rate-of-return, and benefit-cost ratio are the most common methods to evaluate pavement M&R projects (Peterson, 1985).

Although economic analysis is a very useful technique for project selection, it has some difficulties related with the answer to the following questions:

1. Should user cost be included in the analysis? If they are included, they are hard to estimate.
2. What discount rate should be used?
3. What treatments should be applied in the future? When should they be applied? What will be their cost and effects?
4. How salvage value should be calculated?

2.5.2. Prioritization Methods

The ranking criteria used to establish priorities between projects define specific overall goals or objectives of the organization. Similarly to the treatment selection criteria, several methods can be used for ranking projects. Kher (1991) distinguish the following prioritization techniques:

1. **Prioritization by Distress or Performance**. Priority ranking is based in individual distresses or performance characteristics, such as: alligator cracking, rutting of asphalt pavements, roughness, etc.

2. **Prioritization by a Combined Distress-Performance (D-P) Function**. Several distresses or performance measures, such as cracking, rutting and roughness, may be combined in an overall prioritization function or sufficiency rating. This is a fairly common method of ranking.
3. **Prioritization by Composite Criteria**. The next step in complexity, is to combine many characteristics of the road section in the prioritization function (priority rating), such as pavement condition, geometrics, traffic, maintenance history and safety factors.
4. **Prioritization by Least Present Worth Cost**. The projects with lower TPC are given higher priority. It is one of the most appropriate methods since high priority is assigned to those projects that will cost less in the long term.
5. **Prioritization by Benefit/Cost or Cost-Effectiveness**. This method is a very rational way of ranking because higher priorities are assigned to those projects that will return the maximum benefits for the money invested. In the Cost-Effectiveness Method the benefits are measured in subjective terms.

2.5.3. Artificial Intelligence Techniques as Treatment Selection Tools

Treatment selection and prioritization using artificial intelligence techniques, such as expert systems and neural networks can also be considered separate treatment selection and prioritization methods. However, these techniques can also be considered tools that allow to apply one or more of the above methods orderly and efficiently.

2.6. ARTIFICIAL INTELLIGENCE AND PAVEMENT MANAGEMENT

Artificial Intelligence (AI) is a fast growing discipline that attempts to represent and manipulate knowledge to automatically solve problems that before were only solved by humans. It is "the study of how to make computers do things which, at the moment, people do better" (Lawrence, 1994). It includes several technologies such as expert systems, robotics, computer vision, natural language representation, and neural networks.

Because of the nature of the project recommendation, prioritization and selection processes in a PMS, AI techniques can be useful tools that can be applied to these processes. In particular, Expert Systems and Neural Network have been successfully applied in the area of pavement engineering (Alshawi and Cabrera, 1988, Aougab et al., 1989, Attoh-Okine, 1995, Greenstein and Berger, 1987, Haas and Shen, 1987, Hajek et al., 1987, Hajek and Hurdal, 1993, Hall et al., 1987, Harper and Majidzadeh, 1991, Hendrickson and Janson, 1987, Kulkarni, 1985, Lee and Gadiero, 1989, Prahlad et al. 1993, Ritchie, 1987, Ritchie et al., 1987, Ross et al., 1990, Taha, 1995). A general description of both techniques, considering selected applications, is presented following.

2.6.1. Expert Systems

Expert systems (ES) are, "programs, usually confined to a specific field, that attempt to emulate the behavior of human experts" (Schalkof, 1990). They emulate human expertise, and can solve problems which do not have algorithmic solutions. Normally, ES's provide explanation of the reasoning process and conclusions, and may provide useful solutions even with incomplete data (Haas et al., 1994).

Structure of an Expert System

The structure of a typical ES, shown in Figure 2.8, includes the following basic elements (Haas et al., 1994):

1. Knowledge base,
2. Inference engine, and
3. Working memory.

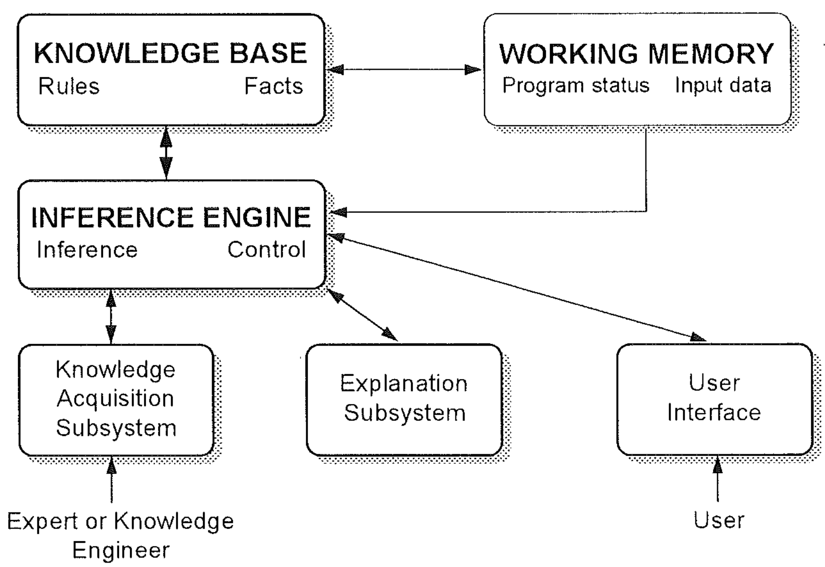


Figure 2.8. Structure of a Typical ES (Haas et al., 1994).

The knowledge base contains the rules, facts and other type of knowledge about the specific problem. The knowledge can be represented using different methods. The most popular approach is the rule based representation. The inference engine, also called rule interpreter, is the responsible of the execution of the program. It receives data from the user about a problem to be solved, and orderly applies the knowledge in the knowledge base to find a solution. The working memory resembles a database of conventional programs. It keeps track of the program status and contains the input data for the given problem.

ES emphasize the need for good user interface and certain explanation modules. For this purpose, in addition to the three basic elements, a typical ES also includes the following components:

1. Knowledge Acquisition/Modification Module, to aid in the translation of the knowledge obtained from the experts.
2. Explanation Subsystem, to explain the reasoning and problem solving strategy to the user, and
3. User Interface, that enables the user to interact with the ES. The user is allowed to request explanations, ask for advice, monitor the decision process, redirect the line of reasoning, etc.

Advantages and Benefits of Expert Systems

Because of the nature of the data used for design, construction, maintenance and pavement evaluation processes, ES technology seems well situated to handle or integrate one or more of these processes. These systems can combine the analysis of quantitative results, expert knowledge, and engineering judgment into an integrated procedure (Haas et al., 1994).

Some of the motivations for developing expert systems are (Schalkof, 1990):

1. "Expert" knowledge is a scarce and expensive resource.
2. ES's make expert knowledge available to a larger audience, making them useful for training, consulting, advising, etc.
3. The integration of the knowledge of several experts may lead to ES's that outperform any single expert.
4. ES's are not motivated to leave companies for better working conditions, to demand high salaries, etc.
5. The development of an ES requires the formalization of knowledge. Explicit knowledge can be replicated, taught, criticized, expanded, etc.

Some of these motivations are very important in the area of pavement engineering where a great deal depends on the judgment and experience of the technologists and engineers. Since many experienced people are nearing retirement and will be replaced by younger people, generally with little experience in pavement technology, it is important to "capture" as much knowledge and experience currently existing (Haas et al., 1994).

It can be argued that PMS's in their current state are rudimentary expert systems. However, PMS's lack a clear division between their inference engines (normally a single decision) and the rule base (typically trigger points for pavement distress severities and extents) and they do not have any explicit explanation capability. ES can be particularly appropriate for pavement management because the rule-base does not remain static, but changes continuously due to the following factors (Flanagan and Halbach, 1987):

1. Because of the continued periodic data collection, new rules can be developed to replace heuristics originally developed by the experts.
2. As the PMS works to improve the road network, the goals of the agency may change.
3. Pavement is a field in which the recognized experts, whose knowledge will originally be incorporated into the knowledge base, have as counterparts local experts, whose experience with local climate, traffic, equipment, materials, work rules, and politics is equally important to the development of a comprehensive knowledge base.

Because all these points, ES can play an important role in organizing and enhancing pavement management practices.

Limitation of Expert Systems

As ES have many advantages and benefits, they also have some drawbacks or limitations, such as:

1. The acquisition and encoding of human knowledge is difficult, slow and expensive.
2. It is difficult to get knowledge of experts because of communication gaps, lack of time, fear for legal/ethical responsibilities, etc.
3. Many experts attribute their success in reaching correct conclusions, not only to the application of a sound reasoning process, but to "gut feel" or "instinct." This implies some kind of subliminal quantification, that is very difficult, if not impossible, to take into account in an ES.
4. Difficult in defining who is an expert in the field.

5. Different experts can recommend different solutions or answer for the same problem.
6. Experts will vary their solutions to a specific problem to gain knowledge by observing the result. Hence, an expert does not always develop a unique solution to a specific problem.

2.6.2. Knowledge Engineering

Knowledge engineering is the process of transforming expert knowledge to a form that a computer can interpret (Lawrence, 1994). The main issues regarding the development of an ES are knowledge representation, knowledge acquisition, knowledge application, and knowledge elucidation or explanation.

Since there are several different schemes for knowledge representation, a key issue in the developing of an ES is to define which one to use. The two main factors that affect this decision are (Barnett et al., 1988):

1. How well the scheme models the given problem domain knowledge, and
2. How easily can problem domain specific inferencing be performed on the knowledge representation scheme.

The most widely used ES knowledge representation scheme is the rule-based scheme. The knowledge is represented using facts and "condition-action" rules. The facts have the form of "object has attribute" or "object 1 -relation- object 2." The rules have the general form "IF -condition is true- THEN -perform action-."

Knowledge acquisition is a critical point in the development of an expert system. It consists of transferring and transforming potential problem-solving expertise from some knowledge source to a problem. The process usually involves a series of lengthy interviews between the group responsible of developing the system and one or more domain experts. If the knowledge about a particular domain is shared by several experts, the process can be conducted though questionnaires and requires the use of specific techniques to quantify the expert's opinions. Some of the techniques that can help in the knowledge acquisition process are briefly discussed in the following sections.

One of the unique characteristics of expert systems is the ability to explain its logic and reasoning to reach the conclusions. This improves the user confidence in the system and allows detection of errors or inconsistencies. The system performance can be evaluated by tracing its reasoning in order to decide whether or not the system is getting the right answer and for the right reasons. Unless a system has a good explanation facility, an expert will not be able to access its general performance or give advice as to how its performance can be improved.

2.6.3. Examples of Pavement Management Applications of Expert Systems.

There have been several successful applications of ES techniques to solve problems in the pavement management area. Some of them are described in the following sections.

SCEPTRE: A Surface Condition Expert for Pavement Rehabilitation

SCEPTRE is a knowledge-based "expert" advising and scooping tool designed to assist highway engineers in planning and developing cost-effective pavement rehabilitation strategies. It was also meant to be used to instruct other engineers, particularly at the local level. It is a rule-based ES that has been developed using the expert system *shell* EXSYS (Ritchie et al., 1987).

The ES recommends feasible M&R strategies for subsequent detailed analysis and design based on project level pavement distress evaluations and other user inputs. The knowledge base was constructed from the extensive experience of two pavement specialists, with extensive experience in pavement rehabilitation in the States of Washington and Texas.

ROSE: A Knowledge-Based Expert System for Routing and Sealing

Rose is an ES developed for the Ministry of Transportation of Ontario (MTO) to advise on recommendations for Routing and Sealing (R&S) of cracks in AC pavements in cold areas. The system recommends section to R&S ordered by desirability, as well as the total amount of cracks recommended for R&S. A rating (1 to 10) indicates the desirability of R&S for a given section (Hajek et al., 1987)

ROSE uses a total of 40 numerical variables and contains about 360 rules. It can operate in interactive or automatic mode. The interactive mode was developed in EXSYS and runs in an IBM compatible computer. It has a user-friendly interface and was used as a prototype to formulate the rules, and to test and calibrate the system. This operation mode allows the user to ask "Why?" data is needed, and to easy review data input. For the automatic mode, the rules were translated to FORTRAN, using an IF...THEN format. The automatic mode interacts directly with the existing PMS database allowing high-speed processing of section and statistical analysis of R&S recommendations.

Researchers found that the development, testing, and calibration of a prototype version of ROSE were made much easier and more efficient by using the inference engine and editing features of EXSYS (Hajek et al., 1987). The interactive version of ROSE required about three months of development and programming (Hajek and Hurdal, 1993).

PARES: A PAvement Rehabilitation Expert System

New Mexico State Highway and Transportation Department (NMSHTD) developed a computerized integrated system to assist in the evaluation and development of rehabilitation schemes for flexible pavements (Ross et al., 1990).

The NMSHTD pavement rehabilitation system uses a priority system, based on field condition surveys and traffic volume, to assess which sections should be rehabilitated. The system has been developed such that a priority assignment indicates that rehabilitation is necessary. Pavements requiring routine maintenance do not receive a priority value. Using the priority assignment, an initial estimate of cost for rehabilitation, and available funds, the system determines the number and extent of projects to be considered for rehabilitation. The factors considered in the selection process are: overall pavement rating value (PMV), individual distress type, severity and extension, ADT, and roughness.

A computerized knowledge-based expert system was designed to allow a more detailed preliminary estimation of rehabilitation needs than with the manual process. The group of state highway experts were used to build the knowledge base. The building process included the construction of a list of the sort of data to be entered by the users into the system and the creation of a hierarchical decision tree. The final logic tree that forms the shell in which knowledge-based rules are manipulated is shown in Figure 2.9.

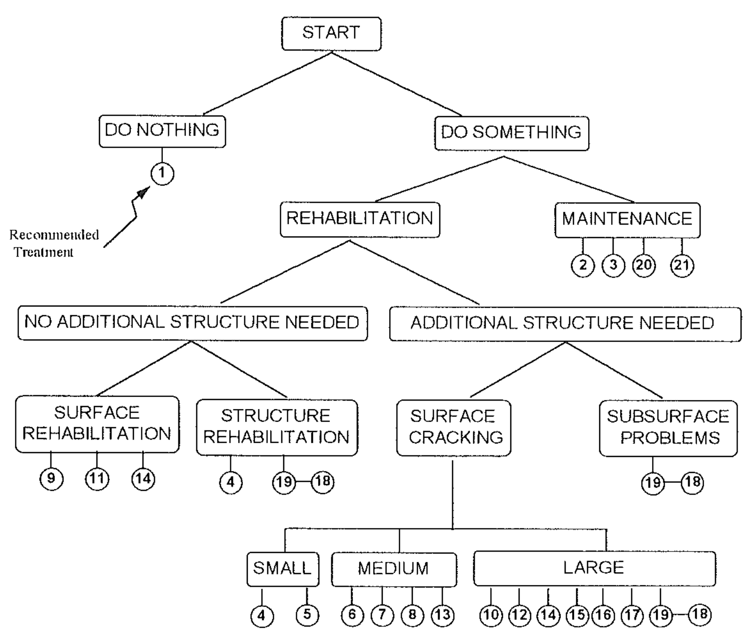


Figure 2.9. Final Multi-level Logic Tree Adopted in PARES (Ross et al., 1990)

The ruled-based Knowledge-Base is composed of 278 rules that can be divided in two main sets or groups. A first set of rules classifies the distresses in five categories. The second set of rules selects appropriate rehabilitation strategies following the multi-lever logic tree. When the execution cycle is complete, PARES ranks the strategies according to how many times each strategy is recommended by a rule (Ross et al., 1990).

Ohio DOT Pavement Management System

The Ohio Department of Transportation (ODOT) constructed an ES that recommends feasible M&R plans to be included in the 6-year plan for each segment within the statewide network. These feasible plans are the input for the optimization process. The optimization selects one of the feasible strategies for each individual roadway section, on the basis of available budget and desired performance (Harper and Majidzadeh, 1991).

The ES develops feasible plans independently for each highway segment using a project level analysis of the individual distresses identified in the section. By selecting a small set of all possible solutions, the ES reduces the solution space for the optimization considerably and has resulted in lessening the computer burden.

The system first screens which M&R would be appropriate for the distress level on each section, for each year of the planning horizon. The future condition of the pavement is predicted using performance models. A second screening further reduces the number of possible treatments using a set of heuristic rules developed in consultation with ODOT design and maintenance engineers. This set of treatments is then optimized to produce the recommended program.

Wisconsin's Pavement Management Decision Support System

Wisconsin Department of Transportation (WisDOT) has developed a rule-based Pavement Management Decision Support System (PMDSS) to provide reasonable and reliable solutions to pavement M&R problems. The collective wisdom of WisDOT engineers and practitioners was encoded as decision rules for problem definition, treatment selection and project prioritization (DeCabooter et al., 1994).

PMDSS uses available data to recommend which type of M&R strategy is needed for a pavement unit but it does not deal with problems of specific pavement design. For example, it will tell that an overlay is needed but, it will not specify the exact thickness. The pavement unit considered are one mile long. The factors considered in the analysis are: severity and extent of individual distresses, Pavement Distress Index

(PDI), Pavement Serviceability Index (PSI), emphasis of the pavement, pavement type, and pavement age. The emphasis level relates with performance expectancy. A high emphasis pavement has higher expectancies and will, in general, receive more intense treatments.

The knowledge base consists of rules for distress evaluation, problem identification and rehabilitation recommendations. PMDSS first assigns an overall severity (minor, moderate or severe) for each distress using the field evaluations. Then, it establishes the nature of the pavement problem and the severity of the problem using the problem's decision elements. After identifying the problems for each section, PMDSS, recommends a range of treatments which would repair all problems in each pavement unit

The pavement units are then combined into project length sections or *improvement sections*. For each section, PMDSS recommends three different treatments, that under-treat 15%, 30% and 50% of the pavement units included in the improvement section. The System then chooses one of the three recommended treatments and prioritizes the selected projects based on the weight the user gives to the various decision five factors plus remaining service lives in terms of PSI and PDI. PMDSS assigns a priority value to each improvement unit.

2.6.4. General Scheme for an Expert System for Project Recommendation

Wang (1992) established the basic conceptual framework for the development of the knowledge-base for an ES for project recommendation for ADOT's PMS. The attribute factors that can be used to determine the most appropriate rehabilitation strategy, as well as the levels for each of them, were tentatively defined. The factors, attributes and levels are shown in Table 2.1.

The objects of this model are the same rehabilitation actions which are used in the decision process by the PMS engineers prior to the pavement engineering design stage. As project selection concerns only with the scope of the preservation required, rather than a specific preservation treatment design, the objects for the ES should be general rehabilitation groups. The same groups used for the Network Optimization System were recommended:

1. Routine Maintenance,
2. Seal Coat, for temporarily sealing cracks;
3. Asphalt Concrete Friction Course or Asphalt Concrete Surface Course (ACSC);
4. Light Structural Overlay, including 2" overlays, ACSC with rubber, etc.;
5. Medium Structural Overlay, including 3" overlay, 2" overlay with rubber, etc.;

Table 2.1. Attributes for Model of Project Selection (adapted from Wang, 1992).

Areas (1)	Factors (2)	Attributes (3)	Levels (3)
Pavement Functional Factors	Ride Quality	Low, Medium, High	3
	Pavement Friction	Good, Bad	2
	Traffic Volume/ADT	0-2000, 2000-10000, >10000	3
Pavement Structural Factors	Percent Crack	Low, Medium, High	3
	Rutting	Low, Medium, High	3
	Subgrade Quality	Good, Medium, Poor	3
Pavement History	Type of Last Rehab. Action	one in the six groups	6
	Time of Last Rehab. Action	0-5 years, 6-10 years, 11-15 years	3
	Surface layer Thickness	4", 5", 6", 7", 8", > 8"	6
	Pavement Alignment	Good, Medium, Poor	3
Environmental Factors	Maintenance Costs/SY	$0.05,$0.08,$0.10,$0.13,$0.15,>$0.15	6
	Drainage	Good, Medium, Poor	3
	Regional Factor	0-1.7, 1.8-2.7, >2.7	3

6. Heavy Structural Overlay, including 4" to 5" overlay.

Once the structure of the knowledge base has been defined, knowledge acquisition should be conducted by interviewing experienced pavement engineers. This process is the most important to ensure the quality and reliability of the knowledge which will be used in the project selection process.

2.6.5. Artificial Neural Networks

Artificial neural networks are models structured upon the organization of a human brain. The human brain is complex network of hundred of millions of neurons that send information back and forth to each other though connections. The result is an intelligence capable of learning, analysis, prediction, and recognition. Artificial neural networks are composed by many simulated neurons or units that are connected in such a way that are able to learn in a manner similar to people (Lawrence, 1994).

Neural networks, also referred to as connectionist systems, have experienced a resurgence of interest in recent years as a paradigm of computational and knowledge representation (Garrett et al., 1990). After a first surge of attempts to simulate the functioning of the human brain using artificial neurons in the

1950s and 1960s, this AI sub-discipline was put on hold until recently. The resurgence has been due mainly to the appearance of faster digital computers which can simulate large networks, and the discovery of new neural network architectures and more powerful learning mechanisms. The new network architectures, for the most part, are not meant to duplicate the operation of the human brain, but rather to receive inspiration from known facts about how the brain works. These architectures are characterized by (Rich and Knight, 1991):

- A large number of very simple neuron-like processing elements, sometimes called artificial neurons, that maintain only one piece of information or level of activation, and are capable of a few simple computations.
- A large number of weighted connections between the elements. The weights on the connections encode the knowledge of a network.
- Highly parallel, distributed control.
- An emphasis on learning internal representations automatically.

Figure 2.10 shows the general scheme for an artificial neural network. At any time, each processor has a specific level of activation, and each connection between two processors possesses a numeric level representing the strength of the connection between them. The neural network is able to store what has learned by changing the strengths of the connection between its processors. A high positive strength of connection from a processor A to a processor B indicates that when processor A is active processor B should also be active. On the other hand, a high negative connection indicates that if A is active B should not be active. A near-zero strength indicates that A should have little effect on B.

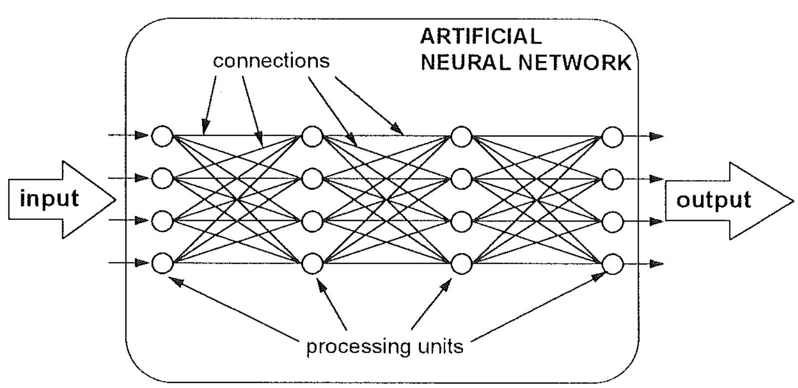


Figure 2.10. Sample Neural Network (after Garrett et al., 1990).

The neural network performs computations by propagating changes in activation between the processors through existing connections. The propagation starts with a "pattern of input," in which a collection of processors receives an excitation or stimuli from an external source (input data). The stimulus is propagated through the network and results in a "pattern of output" in which a collection of processors sends their input to an external receiver that interprets the information (Garrett et al., 1990).

Advantages of Neural Networks

The most important unique characteristics of neural networks that have promoted their development are (Garrett et al., 1990):

1. Neural networks are capable of self-organizing and learning. They are not programmed in the classical sense, but rather they are "trained." A series of examples of the concept to be captured is presented to the system and the network internally organizes itself to be able to reconstruct these examples.
2. They are able to produce correct or near correct responses when presented with partially incorrect or incomplete data. This is particularly important if the information presented in somewhat *fuzzy*.
3. They are able to generalize rules from the cases on which they are trained and apply the rules to new stimuli.

Limitations of Neural Networks

Some of the drawbacks of neural networks, in relation with other forms of knowledge representation, are (Rich and Knight, 1991):

1. Neural networks usually require large number of training examples and long training periods.
2. After a network has learned to perform a difficult task, it is knowledge opaque. It is difficult to retrieve and understand the reasoning developed by the neural network.

Anatomy of a Neural Network

The general structure of neural networks can be divided into seven basic aspects (Rumelhart et al., 1986, Garrett et al., 1990):

1. A set of processors or processing units,
2. The state of activation of a processing unit,
3. The function used to compute output from a processing unit,

4. The pattern of connectivity among the processing units,
5. The propagation rule employed,
6. The activation function employed, and
7. The rule of learning employed.

A neural network is formed by processors or *processing units*, also referred to as nodes, artificial neurons or units. Since there is no centralized control mechanism, all processing is carried out by these artificial neurons. The neurons receive input from its neighbors, modify their state of activation, compute an output, and send that output to their neighbors, as summarized in Figure 2.11. Considering its location within the network structure the processing units can be classified in three types:

1. **Input neurons**, which receive input from external sources and propagates the stimuli to the rest of the network.
2. **Output neurons**, which receives information from the rest of the network, compute their output and sent it to external receivers.
3. **Hidden neurons**, which only connects with other units within the network.

After receiving stimuli, each neuron has an activation level, a_i, which is often represented as a continuous quantity between values of 0 and 1. The state of the network is represented by the activation level of all its processing units.

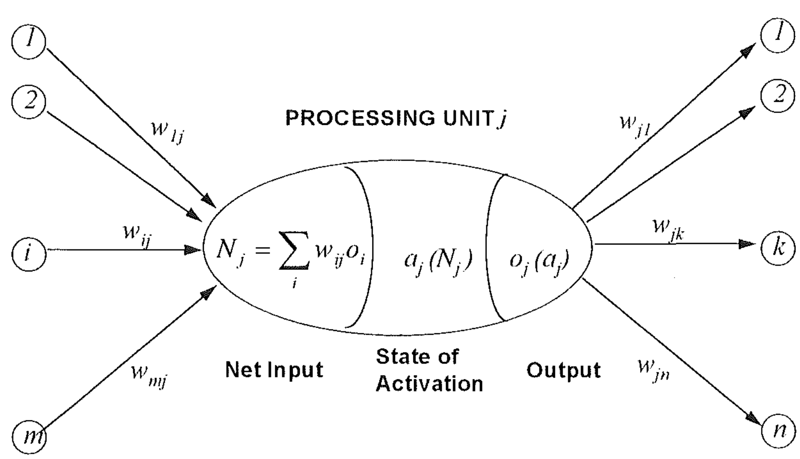


Figure 2.11. Basic Processing Unit.

The *output function*, f_i, is a function that a neuron uses to compute the output it transmits to its neighbors based on its activation level. This output, o_i, is also a scalar value usually between 0 and 1, and its relation with the activation level can be described as follows:

$$o_i = f_i(a_i) \qquad \text{....(2.4)}$$

Some neural networks use a unity function ($o_i = a_i$), and other use threshold function that produces no output unless the activation level exceeds a pre-defined minimum level of activity. This last type of function prevents the network from making "wild guesses" when the input pattern presented does not resemble those that the network has learned.

The *pattern of connectivity* among the processing units and the strength of the connections are the factors that most influence how the network will respond. The connections between two processing units, e.g., from a unit i to a unit j, can be classified according with the weight of the connection, w_{ij} in:

- Excitatory, if w_{ij} is positive and encourages the activation of unit j.
- Inactive, if w_{ij} is zero and has no effect on the activation of unit j.
- Inhibitory, if w_{ij} is negative discouraging the activation of unit j.

The *rule of propagation* describes how the output from other connecting units, o_i, and the strength of the connections, w_{ij}, are combined to compute the net input, N_j, to a processing unit j. The most usual form of this rule is:

$$N_j = \sum_i w_{ij} o_i \qquad \text{....(2.5)}$$

The *activation function* , F, combines the input received by a processing unit j with its current level of activation to compute its new state of activation, and can be expressed mathematically as follows:

$$a_{j-new} = F\,(a_{j-old}, N_j) \qquad \text{....(2.6)}$$

The most simple network simulators use an activation function equal to the value of the net input arriving at the processing unit ($a_{j-new} = N_j$). A more common form for the activation function is a Sigmoid function, that gives an activation level between 0 and 1:

$$a_{j-new} = F(N_j) = \frac{1}{1 + e^{-N_j}} \qquad \text{....(2.7)}$$

The *learning rule* defines how the network is modified in response to the training cases. The most common learning procedure is to modify the strength of the connections between processing units. A very common form of learning rule is the *Delta Rule*, in which the incremental changes in weights are given by

$$\Delta w_{ij} = \eta(t_j - a_j)o_i \qquad \text{....(2.8)}$$

where: η is the learning constant (a value between 0 and 1),

t_j is the expected activation of the processing unit j (given by the solution of the training case), and

a_j is the actual activation level of the processing unit j, and

o_i is the output of the unit i.

Backpropagation Neural Networks

The backpropagation networks are the most popular type of neural network. The term backpropagation is derived from the way the network learns, by propagating the error seen at the output nodes back to the preceding layers of nodes.

The simplest form of backpropagation networks is a two-layer feed-forward network called *Perceptron* that was developed by Rosemblatt (1962). These networks consist of only input units and output units. The input units are activated by the input pattern, they compute their output, and send it to the output units to which they are connected. If the output units receive enough input, they are activated and form the output pattern. Minsky and Papert (1969) showed that these simple two layer networks are incapable of solving a large number of interesting problems. A more common and powerful type of neural networks is the Multi-layer Network, that incorporates one or more layers of hidden units between the input and output layers. These hidden units enable the network to develop more complex internal representations.

Figure 2.12 shows an example of a backpropagation network with only one hidden layer. A regular computation proceeds forward in the network. Given an input pattern that activates the input layer, the output pattern is computed as a result of a forward pass which computes the activity level of each layer based on the activation levels of the layers below it . The neurons in the hidden layer compute their activity levels based on the activation levels of all connected input neurons. The activation of these hidden neurons are used to compute the activation of the output neuron.

The learning process can occur in a supervised or unsupervised fashion. For the supervised learning, the output expected from the network is included in what the network is to learn. On the other hand, unsupervised learning is a more complex process in which the network is not told what it is supposed to learn from the input presented, and must discover by itself regularities and similarities among the input patterns.

Backpropagation neural networks are trained in supervised fashion. They learn by example and repetition. Many example pairs of inputs and output are presented to the network, which adjusts the weights of the connections to minimize the error in the outputs.

The set of examples that is presented to the network, the training set, must embody the function the network is to "learn." The training is an iterative process. Each time the network is presented with an input, the network "guesses" what the output would be. If the network output differs from the one presented in the example, the weights of the connections are adjusted.

Most backpropagation networks use the *generalized delta rule* to adjust these weights (Rumelhart et al., 1986). This rule is a modification of the delta rule presented in equation 2.8.

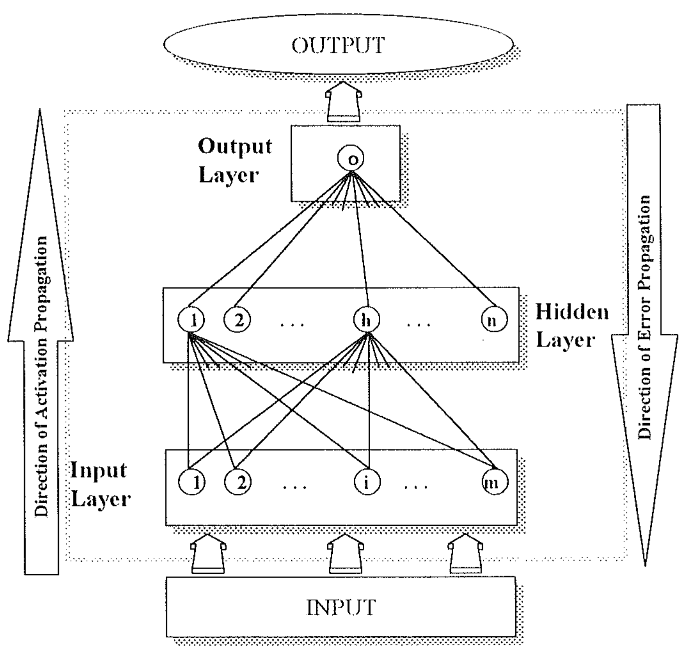


Figure 2.12. Sample Backpropagation Neural Network.

The goal of the learning process is to modify the weights such that the sum of the squares of the errors in the output produced by the network for each input pattern are minimized. The error for a simple output pattern is the sum of the square of difference between the expected activation and the level of activation observed in all output units.

Each training case with the corresponding modifications of connections is called a *cycle*. Each cycle involves three steps:

1. The network is presented an input pattern and propagates the activation forward to the output units (as in a regular computation),
2. The output units then propagates the errors backward to the hidden units, and
3. The units then modify their incoming connections' strengths using the backpropagated error.

A set of cycles, made up of one cycle for each training case, is called an *epoch*. It is common that an artificial neural network requires thousands of epochs to learn most of the examples. The weights are modified many times until (Garrett et al., 1990):

1. The sum of the squares of the errors of all patterns presented during training is below a prescribed value (it has converged on a set of weights),
2. The network gets an acceptable number of right answers, or
3. The number of epoch exceeded a prescribed maximum.

Development of a Backpropagation Network

The development of a backpropagation neural network involves the following steps (Garrett et al., 1990):

1. Verification that a neural network is an appropriate technique to solve the problem that needs to be solved, and that there are a large number of examples available for training. This type of knowledge representation is particularly appropriate for recognizing features within a highly distributed input pattern and uses its information to classify the pattern into one of a set of predefined classes.

2. Selection of the network architecture, implying the selection of the input structure, the number of hidden layers, the number of neurons in each layer, and the interconnectivity between the layers. There are currently no good heuristics for selecting an appropriate architecture. In general, a "good" architecture is one that converges for the cases on which it is trained and performs well for those cases on which it is tested. While the forms of the input and output

layers are dictated by the problem representation, the selection of the form and number of hidden layers is still very much a trial-and-error process.

3. Selection and preparation of the examples on which the network is to be trained and tested. The cases presented to a neural network should embody the various features and sub-features that the network is likely to encounter in a regular working situation.

4. Actual training of the network with the selected examples. The weights are first set at random and successively modified by the iterative process. The amount by which the weights are modified in any given step, or learning rate (η), affects both the speed and oscillation of the convergence.

5. Determination of the ability of the network to *generalize* when presented with testing cases or examples on which it was not explicitly trained. If the network is presented input patterns with features that were not present in any of the training cases, it will probably produce discouraging results. Therefore, it is very important that the training set includes as many features and sub-features as possible. However, because all the features and sub-features, in general, are not apparent to a human, it is usual practice to randomly divide the test cases in a training set and a testing set.

2.6.6. Examples of Pavement Management Applications of Neural Networks.

There have been a few applications of neural networks to solve problems in the pavement management area. Some of them are described in the following sections.

A Neural-Network-Based System for Routing and Sealing

Inspired by the success of ROSE, a knowledge-based ES created to advise on recommendations for Routing and Sealing (R&S) of pavement cracks in cold areas, the Ministry of Transportation of Ontario developed a neural-network-based system for the same purpose. It was hypothesized that the effort to develop and program the ES could be substantially reduced by employing artificial neural network technology. A study was conducted to test this hypothesis and to compare both approaches by evaluating the advantages, disadvantages, strengths and weaknesses of the two approaches (Hajek and Hurdal, 1993).

The problem formulation was the same that for the rule-based ES. The input layer processors were identified with the 40 input factors that were identified as influencing the decision process and the output layer processing units were linked with the desirability values. The neural network was trained with a random sample of 148 pavement sections. The desirabilities for these sections were computed using ROSE and the resulting 148 input/output pairs were used as the training set.

The system was developed using BrainMaker Professional 2.0 software that was considered an appropriate counterpart for EXSYS for the comparison. The supervised training is accomplished using a backpropagation algorithm. The following settings were used to control BrainMaker analysis:

1. A Sigmoid activation function
2. A default learning rate of 1
3. Automatic selection of the number of hidden units
4. Ten percent training tolerance
5. Training facts presented at random

Testing was conducted using 20 additional pavement sections. The solutions approximately agree with the recommendations of the rule based system for high desirabilities (6-10 range). The agreement was not good for lower desirabilities. However, only the rankings of sections with high desirability are important in practice because they are the ones that are scheduled for R&S.

The authors concluded that the neural network solution is considerable easier and fasted to develop than the rule-based system. Neural networks are appropriate for solving selection problems that do not require detailed encoding of causal relationships, for which detailed knowledge is unavailable, or that are not of interest to the users. They would benefit from the development of techniques for interpreting their inner working in terms of causal relationships. The combination of rule-based and neural solutions in one software system or their use for one application could make use the strength of both techniques (Hajek and Hurdal, 1993).

Condition Assessment of Utility Cuts in Flexible Pavements

Prahlad et al. (1993) describe an application of neural networks for predicting a Utility Cut Condition Index (UCCI) to be used as a management tool for identifying conditions of utility cuts and for assigning priorities for their maintenance in the city of Cincinnati, Ohio.

Utilities cut patches are pavement repairs resulting from digging in city streets to inspect or install utility services. The study consisted of a micro level study of localized distresses in asphalt pavements in and around utility cut patches. It started with the development of a Distress Identification Manual for Utility Cuts. Then, the collective judgment of engineers and inspectors was used to establish a UCCI (from

0 to 100) based on the evaluation of a sample of pavement sections using the manual developed, applying the Delphi technique (Prahlad et al., 1993).

A neural network to predict the UCCI was developed using the a commercial neural network simulator, NeuralWorks Professional II/Plus. The original data, consisting of 1,032 observations was separated in two sets. Sixty nine percent of the observations were used for training of the neural network and the remaining 31 percent was used for testing. Thirty processing units were used in the input layer to represent the extent and severity of nine types of distress in the utility and six in the vicinity. The hidden layer consisted of 10 processing units. The output layer had only one unit to represent the UCCI for each utility cut. The network used a Sigmoid activation function and learning coefficients of 0.3 for the hidden layer and 0.2 for the output layer. These factors were obtained by trial and error.

A comparison of actual UCCI, with respect to predicted UCCI, for the testing data showed an absolute average error of 6.5 distress units and an average relative error of 4 percent. Ninety-two percent of the output was included in a ± 12 distress units' band. Most of the discrepancies occurred when UCCI was larger than 90 or lower than 10.

Given the large variation in the condition data used in the training stage, these results were considered satisfactory.

2.6.7. Expert Systems vs. Artificial Neural Networks

Expert systems and artificial neural networks are two different artificial intelligence approaches, each with distinct strengths and weaknesses. Expert systems are rule-based and are best suited for solving problems that follow sequential logic. Neural networks depend on examples and are most appropriate to solve problems involving pattern recognition. A problem solving "specialist" can be produced from either technique, by training a neural network or engineering an expert system (Lawrence, 1994).

In general, expert systems are easier to understand but more difficult to implement. While the process of designing of an expert system often takes months, a neural network may be designed and trained in days, after the examples are gathered. Building the rules for the knowledge based of an expert system is usually difficult and time consuming. However, it can have positive effects, allowing verification of the rules, identification of discrepancies, and development of precise guidelines. The existence of rules may allow other experts to supplement or correct the knowledge base, and it can result in a system that out performs any single expert.

An advantage of expert systems is that they are capable of explaining their conclusions. Expert systems are excellent for training and standardizing complex procedures. In contrast, neural networks may appear to the user as a "black box" with a limited ability to explain itself. However, in many cases when the user does not need to know all the steps in the decision-making process, neural network may be the most appropriate solution.

The type of data available is a critical factor in deciding between a neural network and an expert system. Neural networks are particularly useful when the rules are not clearly defined, or there is no interest in generating them, but there are a lot of examples with relevant data. An important strength of neural networks is that they can handle some amount of inaccuracy, uncertainty or contradictory data, if there are enough examples. This allows them to have greater generalization ability. In addition, updating of neural networks by increasing the training set, is quite efficient and simple compared with updating rule-based systems. If no examples are available, but there are clearly stated rules and formulas, an expert system will probably work better.

2.7. TECHNIQUES FOR GATHERING AND QUANTIFYING EXPERT'S OPINIONS

Several approaches were considered for developing a pavement preservation prioritization formula based on expert's opinions. The main techniques for gathering and quantifying expert opinions analyzed were:

1. Utility Theory,
2. Bayesian Approach,
3. Delphi Technique,
4. Rational Factorial, and
5. Fuzzy Sets Theory.

2.7.1. Utility Theory

Utility theory takes into account the relative magnitude of payoffs and their expected value in the decision making process (Jones, 1977). Basically, the procedure involves the assessment of scalar utility functions which express a decision maker's preference over different level of selected attributes. It assumes mutual preferential independence between attributes.

The process of evaluating alternatives in terms of overall desirability, when there are several attributes to consider, involves the following steps (Finn et al., 1976):

1. Establishing scalar utility functions for each attribute incorporating the decision maker's attitude toward risk,
2. Assessing tradeoffs between conflicting attributes, and
3. Calculating expected utility, combination of the utility functions for each attribute, for each alternative strategy.

The utility functions and tradeoffs can be obtained by interviewing a group of experts in the field under consideration (Finn et al., 1976). Utility theory has been used to develop overall pavement performance indexes in Arizona (Finn et al., 1976) and Texas (Hudson et al., 1982).

2.7.2. Bayesian Approach

The Bayesian Methodology can be used to quantify and combine subjectively and objectively obtained data to develop a prediction model. It is possible to use the past experience of highway engineers in a statistical format similar to that used to analyze data obtained from experiments. New experiences gained from field performance can be combined with previous experience, or *prior* information, to arrive at what is called *posterior* information to develop a predictive equation (Smith et al., 1979).

The process begins with the definition of the variables that will be incorporated in the model. Then, the objective and subjective information regarding the response for the different levels of the selected variables must be collected. The information is processed and two separate statistical models are developed for the objective and subjective data respectively. The models are then combined using the *Bayes' Theorem*. This theorem states that the posterior probability of an unknown state of natura, θ, is given by:

$$\text{Posterior } P(\theta)_{\text{given a sample output}} = C\ (\text{Sample Likelihood}_{\text{given } \theta})\ (\text{Prior } P(\theta)) \qquad \text{....(2.9)}$$

Where $P(\theta)$ = probability of θ

C = normalizing constant

Smith et al. (1979) presented an application of this methodology to develop a model to predict the occurrence of fatigue cracking in asphaltic pavements as a function of four controlled design variables. Expert opinions and measured data were combined to quantify the mathematical models. The research used Bayesian regression analysis, in which the regression parameters are assumed to be random variables with associated probability distributions.

2.7.3. Rational Factorial

In the Rational Factorial Methodology the experts' opinion regarding a particular matter is obtained using a designed "experiment." The process starts with a definition of the factors and levels of the items that influence the topic under investigation. These factors and level are used to define a factorial experiment. If the number of factors is large, a fractional factorial experiment design can be used. This type of design allows identification of the significant variables in the study as well as direct estimations of the relative effect of each independent variable. Depending on the type of fractional factorial selected, some interactions between the independent variables can also be estimated (Zaniewski et al., 1984).

A questionnaire is then designed and tested to obtain the opinion of the experts participating in the study. The questionnaire must include detailed filling instructions and clear definitions of the various factors and levels considered. The questionnaire is distributed, filled out, returned to the researches, and checked for completeness. The results can be analyzed using regression analysis.

This methodology has been used to develop priority indices of pavement performance for the state of Texas (Fernando and Hudson, 1983) and at the national level (Zaniewski et al., 1984). The national study considered seven factors, pavement type, road class, Pavement Condition Index (PCI), PSI, traffic level, temperature (freeze-thaw), and rainfall. Figure 2.13 shows the one quarter factorial experiment used for the study. This design allowed estimation of all main factors and some two-factor interactions.

Pavement Type			Rigid								Flexible							
Road Class			Interstate				Primary				Interstate				Primary			
PCI			20		80		20		80		20		80		20		80	
PSI			2.4	3.5	2.4	3.5	2.4	3.5	2.4	3.5	2.4	3.5	2.4	3.5	2.4	3.5	2.4	3.5
Traffic																		
Wet	F/T	High	x							x				x	x			
	no F/T	Low			x			x				x					x	
	F/T	High		x			x				x					x		
	no F/T	Low				x	x				x							x
Dry	F/T	High					x						x			x		
	no F/T	Low				x	x				x					x		
	F/T	High	x							x				x	x			
	F/T	High			x			x			x						x	

Figure 2.13. Example of Fractional Factorial Design (Zaniewski et al., 1984).

This figure indicates the 32 combination of factors, that were presented to the experts in the survey questionnaire. These experts were asked to indicate the priority they would assign to each sections and if the section should receive M&R treatment. The equation obtained showed that the factors that are significant for the priority index were PCI, PSI, traffic, freeze thaw, rainfall, and PCI/PSI interaction.

2.7.4. Delphi Technique

The Delphi technique attempts to achieve consensus among a group of experts, through cycles of intensive questioning interspersed with controlled opinion feedback. It consists of an iterative procedure characterized by three features:

1. Anonymity,
2. Iteration with controlled feedback, and
3. Statistical response.

This technique avoids direct confrontation of experts with each other, avoiding difficulties inherent with face-to-face interaction, such as spurious influence of high status individuals, ego commitments, and group pressure for conformity. Each panelist must fill out a survey. After the survey is completed, feedback is provided regarding the summary of the results. If there are wide variations in opinions, the panelists are asked to revise their answers until a reasonable agreement is reached, or it becomes evident that such agreement cannot be reached.

2.7.5. Fuzzy Sets Theory

Fuzzy sets, introduced by Zadeh (1965), are used to transform ambiguous and fuzzy or vague information into numerical data in a systematic way. This allows the use of subjective information, such as expert opinions, rules of thumb, and other "non quantifiable" but significant information, in the decision process.

Fuzzy systems allow the processing of vague information in a logical fashion. They use fuzzy numbers or sets and fuzzy logic. They work with "degrees of truth" or "degrees of belonging" to one class or another (Lawrence, 1994).

Representation of Fuzzy Information

An ambiguous or fuzzy input, usually given in linguistic terms such as short, long, good, poor, bad, etc., can be translated into a fuzzy set (fuzzy number) rather than into a single number. Instead of using a simple number to quantify certain element, a "membership" function, $f(x)$, that indicates the "degree of belonging" to each level of the domain element x is used. Figure 2.14 shows four classes of fuzzy numbers:

- Class I fuzzy numbers are used to represent a point estimate indicated as "about *m*."
- Class II fuzzy numbers are used to represent an interval estimate
- Class III represents the notion of "greater than about *m*," and
- Class IV fuzzy numbers are used to represent the notion of "less than about *m*,"

Table 2.2 shows an example of representation for a three-level subjective estimation of the plasticity index of a soil using the fuzzy numbers classes shown in Figure 2.14.

Processing of Fuzzy Information

Special processing methods are required to process fuzzy data. Fuzzy logic operations are similar to traditional logic operations, but very different from probability operations. With probabilities an AND is processed by multiplying the variables, or by adding the values of the variables. In fuzzy logic, ANDs and ORs are processed by taking the minimum or maximum values respectively. In a fuzzy set the minimii and maximii are processed pairwise. The intersection of two sets (AND) is found by picking the smallest of the two elements in each pair. (Lawrence, 1994).

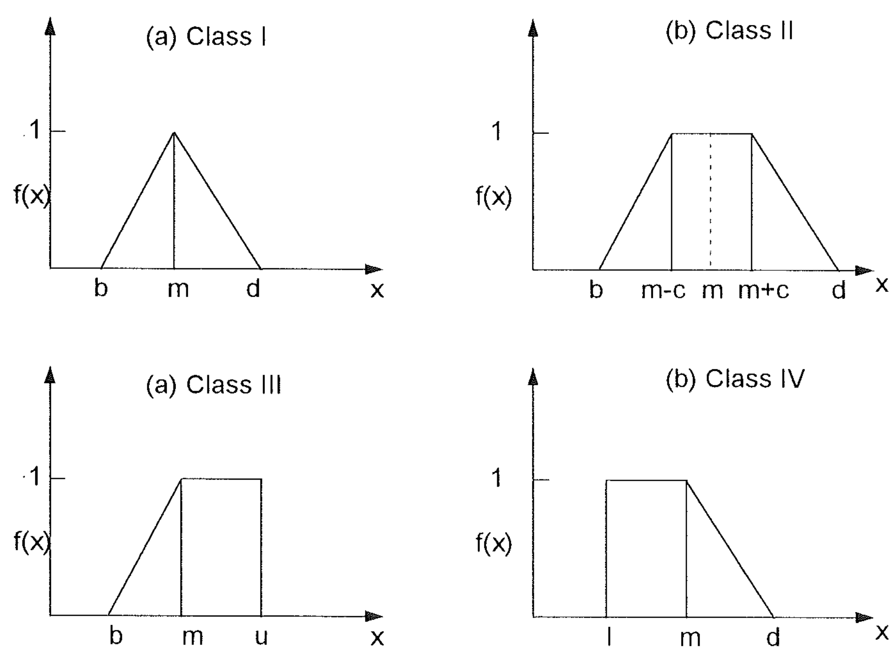


Figure 2.14. Four Classes of Fuzzy Numbers (Juang et al., 1993).

Table 2.2. Example of Linguistic Terms and Their Corresponding Fuzzy Numbers for Representing the Plasticity Index of a Soil (Juang et al., 1993).

Linguistic Term for Describing PI (1)	Fuzzy Number Characteristics						
	b (2)	m (3)	c (4)	d (5)	l (6)	u (7)	Class (8)
High	25	30	--	--	--	50	III
Medium	10	20	5	30	--	--	II
Low	--	10	--	15	0	--	IV

Interpretation of Fuzzy Output

Since the output of a fuzzy problem solving approach is a fuzzy number, rather than a single number, it is necessary to translate it back to a linguistic term that describe the output. The translation requires three elements:

1. A group of standard linguistic terms commonly used to describe the subject matter, such as good, fair, and poor.
2. A group of fuzzy sets, each of which represents one of the standard linguistic terms.
3. A model for determining similarity between the output fuzzy set and each of the standard terms.

The most appropriate translation is the linguistic term whose fuzzy set is closest to the output fuzzy set.

Fuzziness in Neural Networks

Artificial neural networks can employ fuzzy principles in several ways (Lawrence, 1994):

1. They can handle fuzzy data,
2. Can interface to a fuzzy system, or
3. Can include fuzzy principles in neural network algorithm itself.

The first option is the most common. It is not rare, that neural networks use somewhat imprecise input data that can be quantified using fuzzy sets. In addition, the output of an artificial neural network can be also fuzzy and may need to be translated into an appropriate recommendation. For example, a neural network trained to recommend pavement preservation treatments, may give an output that is 0.2 seal coat, 0.6 thin overlay and 0.4 medium overlay. This result has to be "defuzzified" into a specific treatment for the final recommendation. This recommendation could be a thin overlay.

Fuzziness in Expert Systems

Since it is rare that a problem can be adequately described in terms of black and white, or yes or no, the ability to handle fuzzy data and rules enhance the capabilities of an expert system. A fuzzy expert system may use fuzzy data or output but apply probabilistic rule, or may apply fuzzy logic in their reasoning. Since probability measurements cannot take fuzzy sets as arguments, the first type of fuzzy expert system requires three main steps: defuzzification, rule evaluation, and fuzzification. The system takes the fuzzy data and membership functions and translates them into non-fuzzy data, applies the rules, converts the result again to fuzzy sets, and translates them to a meaningful output.

Application of Fuzzy Set Theory to the Pavement Management Area

Pavement management decisions usually involve extracting numerical inputs from subjective information. This process generally induces uncertainties. If these uncertainties are of an ambiguous rather than a random nature, the use of fuzzy set theory rather than probabilistic theory is recommended (Juang et al., 1993). Fuzzy sets theory has been used for several specific applications in the pavement management area (Elton and Juang, 1988, Juang and Amirkhanian, 1992, Zhang et al., 1993).

Elton and Juang (1988) reported a proposed method of asphalt pavement evaluation using fuzzy sets that attempt to reduce the need of an expert to perform the pavement evaluation. The method used the experience of past experts and fuzzy arithmetic. A computer program that combines the evaluation of five individual distresses into an overall pavement rating was developed. The program uses fuzzy arithmetic to combine the distresses and produce a fuzzy set representing the pavement condition that is translated to a linguistic term that qualify the pavement.

Zhang et al. (1993) presented a methodology to develop an index model, called the Overall Acceptability Index (OAI), for flexible pavement using fuzzy sets concepts. The OAI combines four basic attributes or evaluation measures, roughness, surface distress, structural capacity and skid resistance, and their relative importance into a single index that gives a comprehensive evaluation of a pavement's condition.

2.8. SUMMARY AND CONCLUSIONS OF THE LITERATURE REVIEW

This section presented a comprehensive literature review of many aspects related with the different project recommendation, selection and prioritization procedures that can be used in a project selection level PMS.

A M&R program is composed of projects to be rehabilitated and specifies when and how to rehabilitate each project. Programming is the matching of available projects with available funds to accomplish the goals of a given period.

The first step in the project selection process is to delineate roadway project units or sections with approximately homogenous physical characteristics and condition. Each uniform or homogeneous unit is a candidate project. There are several methods of unit delineation. The best solution appears to be a delineation procedure that combines historical data and pavement condition to determine the final section delineation.

The capability of modeling pavement performance is an important feature of both project and network level pavement management. It allows the PMS to use more sophisticated analysis techniques, such as remaining service life and life-cycle cost analysis, and to produce multi year M&R programs.

The remaining service life of a pavement section is a very important pavement performance indicator that should be considered in the project selection process. It can be useful as an indicator of current and future pavement rehabilitation needs, to assess the impact of various budget levels on the health of the network, and to plan yearly balanced programs.

The criteria for treatment selection and project prioritization depend on the agency requirements, availability of resources and sophistication desired. They must be compatible with the specific overall goals, or objectives, of the organization.

Artificial Intelligence techniques can be used to enhance the project recommendation, prioritization and selection processes in a PMS. Both expert systems and artificial neural networks have been successfully applied in the area of pavement engineering. A general description of both techniques, considering selected applications, has been presented.

Expert systems are programs, usually confined to a specific field, that attempt to emulate the behavior of human experts. Expert systems can solve problems which do not have algorithmic solutions, having several advantages with respect to conventional software.

The acquisition and encoding of human knowledge is difficult, slow and expensive. It is difficult to defining who is an expert in the field and hard to get knowledge from experts. Several techniques, such as utility theory, Bayesian approach, rational factorial, Delphi technique and fuzzy sets modeling can be used to facilitate the knowledge acquisition process.

Artificial neural networks are structures of many simulated neurons or units, structured upon the organization of a human brain, that are connected in such a way that are able to learn in a similar manner to people. The most important unique characteristics of neural networks are that they are capable of self-organizing and learning, they are able to produce correct or near correct responses when presented with partially incorrect or incomplete data, and they are able to generalize rules from the cases on which they are trained and apply them to new sets of data.

Both AI techniques appear to be appropriate to use in a project recommendation procedure for a PMS. However, they are two different artificial intelligence approaches, each with distinctive strengths and weaknesses. Expert systems have the advantage of providing better explanation capabilities and the need to develop the rules for a rule-base expert systems allows verification of the rules, identification of discrepancies, and development of precise guidelines. Expert systems are best suited for solving problems that follow sequential logic and are excellent for training and standardizing complex procedures. On the other hand, neural networks do not require the development of knowledge rules, have greater generalization capabilities, are faster to develop and allow easier updating of the knowledge base. Neural networks are particularly appropriate to solve problems in which rules are not clearly defined, or there is no interest in generating them but there are a lot of examples with relevant data.

3. PROJECT DELINEATION

Any pavement preservation project selection process requires the roadway network to be partitioned into roadway units or management sections with approximately uniform physical characteristics and condition.

It is normally more cost-effective to rehabilitate projects that are homogenous in condition than to rehabilitate projects that do not have uniform conditions (Kuo et al., 1992). If the project has sub-sections with different condition, and a preservation treatment is recommended based on the average condition, some of the sub-sections could be over-designed and others under-designed. Therefore, homogeneous management units allow a more effective and efficient programming.

This section describes an automatic project delineation routine developed and implemented for the comprehensive Project Recommendation Procedure. This routine starts the programming process. It divides the pavement network in uniform management units and computes average physical characteristics and condition indicators.

3.1. GENERAL REQUIREMENTS

The first step in the development of the project recommendation procedure was to define a section delineation methodology. Although the delineation can be conducted manually, an automatic delineation procedure has several advantages:

1. It reduces the time and effort required to determine the uniform sections. Since ADOT has more than 12,000 kilometers (7,500 miles) of pavements, the partitioning of the network by hand would require a considerable amount of time and effort.
2. It is expected to reduce subjectivity.
3. The sections limits can be easily revised each year, reflecting the most recent changes in roadway conditions.

Consequently, the methodology for the section delineation was selected taking into account the possibility for automatization.

3.2. SELECTED PROJECT DELINEATION METHODOLOGY

As discussed in Section 2, the uniform units or sections should have not only the same pavement type, construction history, and traffic, but also relatively low variability of the response variables, such as serviceability index and surface condition along the section. For this reason, the selected delineation procedure combines an idealized approach, based in pavement characteristics and historical data, with measured pavement condition to determine the final section delineation, as shown in Figure 3.1.

The delineation procedure must verify minimum and maximum section length constraints. These constraints were defined in conjunction with ADOT PMS engineer.

3.2.1. Delineation Using Historical Project Information

An automatic delineation procedure requires that all information is located in the PMS databases. From the available information, the following roadway characteristics were selected to determine section boundaries:

1. Roadway classification,
2. Roadway denomination,
3. Last construction or rehabilitation project, and
4. Maintenance district.

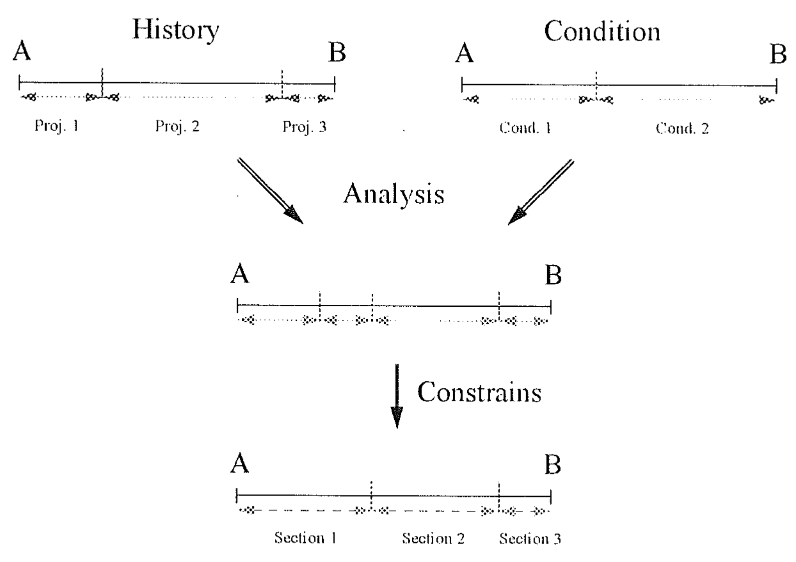


Figure 3.1. Scheme of the Selected Section Delineation Procedure.

Two options where considered to determine which past pavement construction and rehabilitation records to use:

1. All historical project information, or
2. Only the last project applied to each milepost.

The first alternative delineated sections using all project records available in the projects database (PROJMOD.DBF). This option resulted in a large number very short sections. Since most of the sections were shorter than the desired minimum length constraint, this alternative was not acceptable.

The delineation procedure that divided the sections using only the information of the last project was implemented. This information is an element of the main PMS database (PMS.DBF). However during this research it was discovered that this information was not current. A utility computer program was prepared to update the last project information in the main database using the information in the projects database (PROJMOD.DBF). The program updated construction dates, pavement lifts and thickness, and computed the modified structural number. This program was incorporated to the software package into allow future updates.

3.2.2. Delineation Based on Pavement Condition

The unit delineation by cumulative difference method described in section 2.2.2 was selected to delineate statistically homogeneous units from the pavement condition measurements. This method was selected because it is a relatively straightforward and powerful method that gives reasonable results. It is recommended by the AASHTO Guide for Design of Pavement Structures (AASHTO, 1993). This method also has the advantage of being readily adaptable to a computerized solution. The delineation procedure uses the cumulative difference, Z_x, defined as the difference between the area under the cumulative response curve at any distance and the area developed by the overall project average response at the same distance. When Z_x is plotted as a function of distance along the project, unit boundaries occur at the locations where the slope (Z_x versus x) change sign. These points correspond to local maximums and minimums of the Z_x curve.

Several variables, such as surface distresses, roughness, skid resistance, traffic, and maintenance cost, were considered for the delineation phase based. Roughness and cracking were selected at this stage. More variables can be easily added in the future if it is considered necessary. However, trial delineation attempts using more variables resulted in excessive number of sections.

The process was slightly modified when using roughness as the response variable. The original procedure computes the average response for a given interval as the average of the measurements at the extremes of the interval, as indicated in equation 2.3. However, this step was not necessary in the case of roughness because the PMS database already contains the total roughness between mileposts normalized to 1.6 km (1 mile) intervals. For example, the value indicated for milepost 1 is the total roughness between mileposts 1 and 2.

The regular procedure was followed for cracking. The average cracking for each roadway mile was computed as the average of the measurements taken at the milepost indicated and at the following milepost.

In addition, the Z_x plots for real field-measured response variables are very irregular. They show many small oscillations that would result in excessive section boundaries if considered as local maximums and minimums. To avoid this problem, the curve was "smoothened." The implemented delineation procedure considers that a point represents a section break only if the parameter Z_x is lower (or higher) than the two previous and the two following values.

Several options were studied and the recommended section breaks were compared with those determined by visual observation of the condition plots along the route lengths. The selected "smoothing rule" was the one that resulted in the set of section boundaries closer to the boundaries determined manually.

3.2.3. Analysis Sequence

Two alternatives of how to combine historical and condition information for project delineation were evaluated:

1. The first alternative studied the cumulative difference procedure considering the entire length of a particular route, e.g., Interstate 8 in the eastbound direction, and then superimposed the boundaries obtained with those from the construction history analysis, and
2. The second alternative applied the cumulative difference procedure to the sub-sections determined using the last project information. The sections were considered for further partitioning only if they were longer than a minimum section length.

Both criteria were used to compute all boundaries for several complete routes (approximately 1,000 km or 625 mi.). The section boundaries obtained in each case were compared with those established by the researcher based on visual observation of the plots of the pavement condition indicators considered along the route lengths. As an example, Figure 3.2 shows the comparison of section boundaries for Interstate 8 in the eastbound direction. The figure shows the plots for roughness and cracking, as well as the boundaries determined using:

1. Historical project information only (History),
2. The cumulative differences procedure for roughness and cracking considering the entire route (Alt. Con.),
3. The cumulative differences procedure for roughness and cracking within each section determined based on historical project data (Condition),
4. The adopted procedure, that combines 2 and 3 (Adopted), and
5. Visual observation of the pavement condition plots (Manual).

The plots for the other sample routes studies are provided in Appendix I. The alternative that applies the cumulative differences to subsections determined using the last project information was selected for implementation because the sections obtained were closer to those obtained by visual examination of the pavement condition plots.

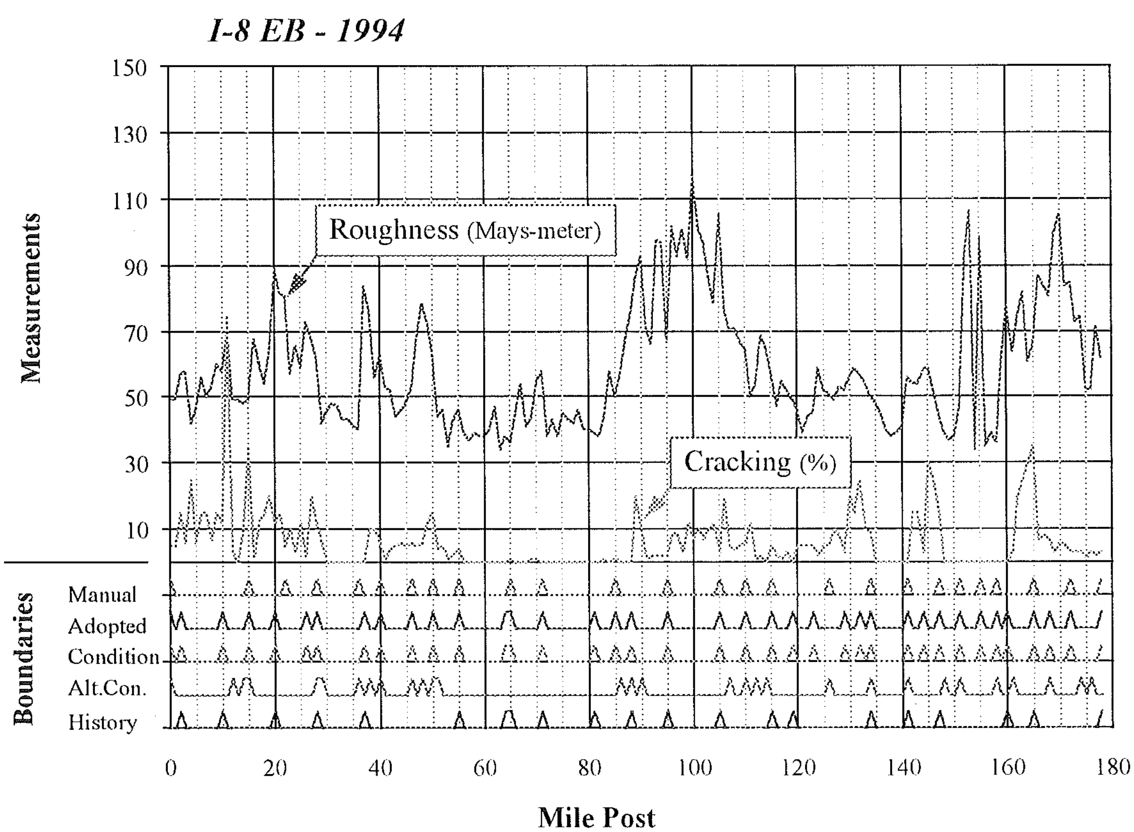


Figure 3.2. Comparison of Alternative Delineation Methods.

For the alternative where the cumulative differences procedure is applied to the entire route, Z_X is computed using the overall project average condition. This alternative identified some section boundaries that appeared unreasonable if compared with those obtained by visual examination of the pavement condition plots. These unreasonable section boundaries defined by this alternative may be explained by the fact that the section delineation decision is then influenced by the condition several hundreds of miles away from the point under consideration.

3.2.4. Additional Considerations

The section breaks determined using the cumulative differences were incorporated only if subsequent sub-sections had average condition significantly different between each other. The program uses default criteria for determining if the difference in condition is significant enough to warrant a section breakpoint. A statistical t-test comparison of means was attempted but did not work. In most of the sections the standard deviation was too large and some of the sections have too few points to even compute a meaningful standard deviation.

The default criteria were established by comparing the boundaries obtained for the testing routes with those determined by visual observation of the condition plots. The criteria can be modified by the PMS users. The default criteria used to decide if two subsections have significant different responses are the following:

1. A section boundary due to variations in roughness is kept if the average roughness difference for the resulting subsections exceeds:
   - 10 Maysmeter units, and
   - 25% of the average roughness of the lower-roughness subsection.
2. Similarly, a section boundary due to variations in percentage of cracking is kept if the average cracking difference of the resulting subsections exceeds:
   - 3% of the pavement surface, and
   - 50% of the average cracking percentage of the lower-cracking subsection.

The tolerances for cracking are higher than for roughness, because the cracking measurements show a much higher within-section variation and cracking is measured only in a small sample (90 m^2 or 1,000 ft^2) of each 1.6 km (1 mile) database unit.

3.3. IMPLEMENTATION OF THE PROJECT DELINEATION PROCEDURE

The project delineation methodology was implemented in a computer program written in Microsoft® FoxPro® 2.6 for Windows. The program automatically determines sections by combining unit boundaries determined using historical data and measured condition. It identifies and uses the most recent condition data available (last year input). The general scheme for the delineation process is presented in Figure 3.3.

The steps conducted by the computer program are listed following:

1. Establish sections breaks based on the last construction project as indicated in the main PMS database. Projects only one mile long where combined with their closest neighbors if the last rehabilitation dates where less than two years apart. The two-year limit was defined as a default and can be changed by the user.

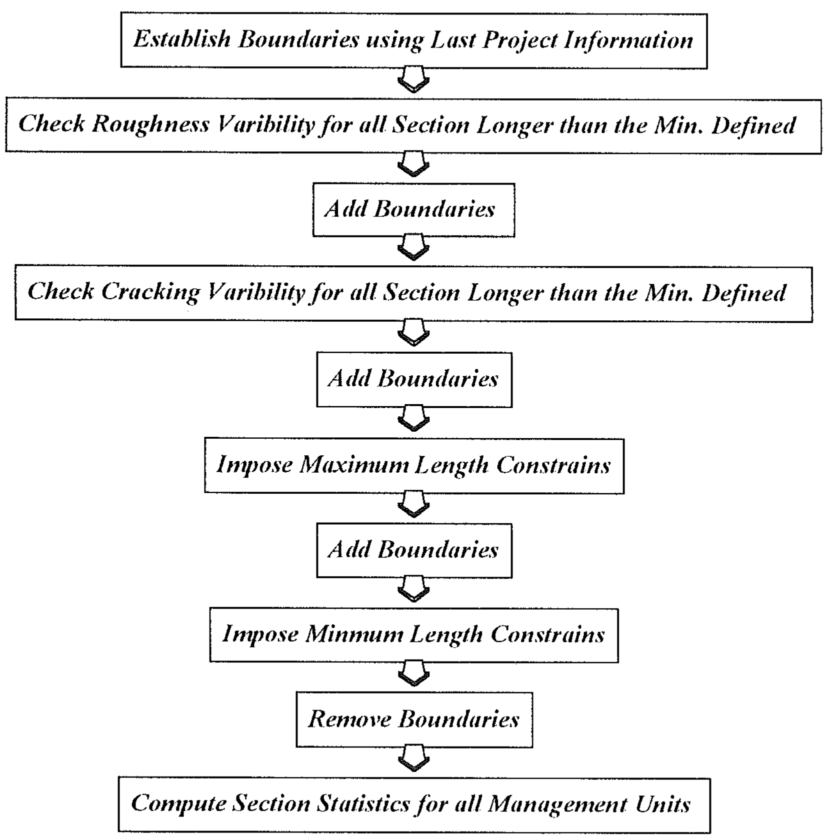


Figure 3.3. General Scheme of the Section Delineation Procedure.

2. Apply the cumulative difference procedure to each section longer than the threshold value or minimum project length and determine if it is necessary to divide the section due to significant differences in roughness between sub-sections. The minimum project length was established to avoid dividing sections that are already of a reasonable length. The default minimum project length used is 8 km (5 mi.), but this can be modified by the user. Establish new section boundaries if necessary due to significant differences in roughness between sub-sections.
3. Reapply the cumulative difference procedure to each section longer than the minimum length, but now using cracking as the pavement condition indicator. Establish new section boundaries, if it is necessary due to significant differences in cracking between sub-sections.
4. Impose maximum length constrains. The default maximum length is 19 km (12 mi.) for interstate and 38 km (24 mi.) for non-interstate highways, but these can also be modified by the user.
5. Review section breaks and impose minimum length constrains, 8 km (5 mi.) by default, only if contiguous projects have reasonably similar condition and construction or rehabilitation dates. If they have similar average characteristics, group them into one section by eliminating the section break.
6. Once all section breaks has been verified, compute average section characteristics and pavement condition indicators for each uniform management unit.

The maximum differences in pavement condition allowed to combine two adjacent sections while imposing minimum length constrains (step 7) are shown in Table 3.1. These values are only default guidelines that can be modified by the user.

Table 3.1. Section Comparison Threshold Values.

Pavement Characteristic (1)	Allowed Percent. of Variation (2)	Allowed Maximum Difference (3)
Region	25 %	0.10
Structural Number	25 %	0.50
Cracking	50 %	3 %
Roughness	25 %	10 Mays-meter Units
Maint. Cost	50 %	$ 300
ADT	25 %	500 vehicles
Rate	25 %	5 points

3.4. ANALYSIS OF THE UNIFORM SECTIONS DEFINED

The unit delineation program was run using the 1994 pavement condition data. As a result, ADOT's pavement network of 12,000 kilometers (7,500 miles) was divided in 1,765 sections.

A distribution of section lengths is shown in Figure 3.4. The average section length was approximately 6.8 kilometers (4.3 miles). The number of sections obtained is relatively large and many projects are shorter than the desired minimum length. This is due mainly to considerations regarding the construction and rehabilitation history of the sections. The database records many short projects and in many cases subsequent treatments over a specific roadway section have slightly different starting and ending mileposts. These differences create small sections with different history.

Combining projects that have been constructed or rehabilitated at different times, or with significantly different condition, would make it impossible to estimate the remaining service life of the sections as is required in the next step of the Project Recommendation Procedure.

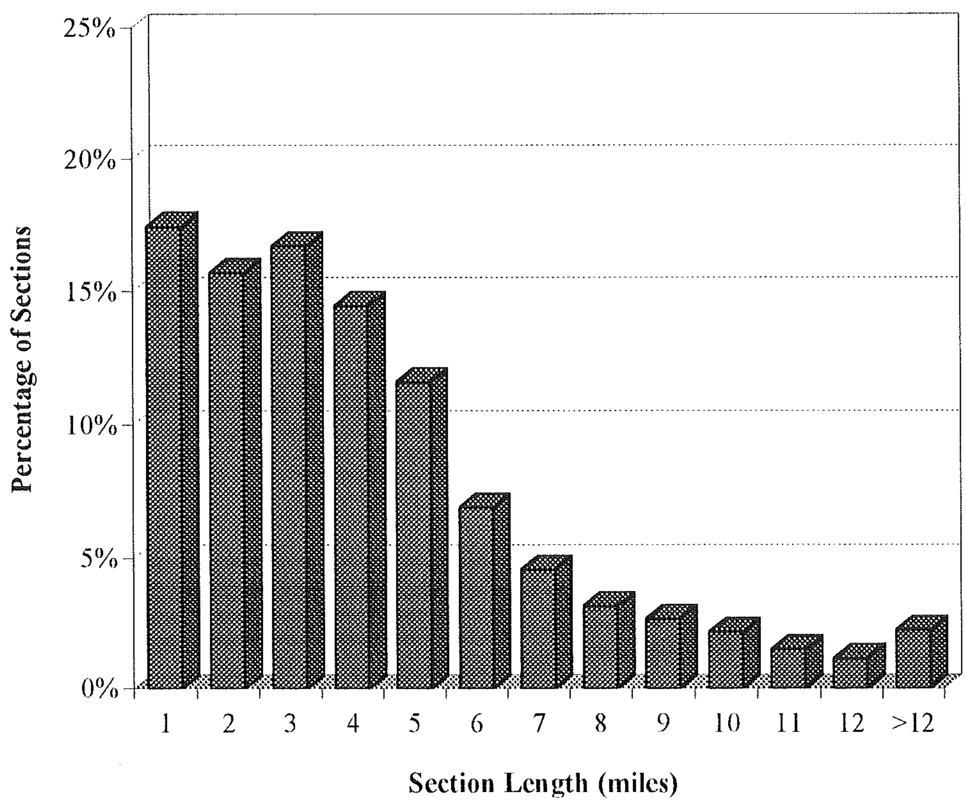


Figure 3.4. Distribution of Uniform Management Unit Lengths.