Optimal Maintenance Decisions for Pavement Management

By J. V. Carnahan,[1] M. ASCE, W. J. Davis,[2] M. Y. Shahin,[3] P. L. Keane,[4] and M. I. Wu[5]

(Reviewed by the Highway Division)

Abstract: The object of this investigation was to develop a procedure for making optimal maintenance decisions for a deteriorating system. The particular system chosen for study is a pavement, and a methodology is developed to ensure that pavements meet certain performance criteria while minimizing the expected maintenance cost. A cumulative damage model based upon a Markov process was developed to model pavement deterioration. The optimal repair action for each possible pavement state in the planning horizon was found by means of probabilistic dynamic programming. Sample sequences of repair actions were generated during a simulation in which the optimal repair policy was applied to sample pavement condition histories. Several sensitivity studies were performed to study the variation in expected cost, including the effect of delaying the optimal program.

Introduction

The maintenance of deteriorating facilities can be a great challenge because the process of deterioration has random features for some facilities. For example, the effect of weather and traffic loads on pavement life is nearly unpredictable. Despite such uncertainties, typical questions that may arise regarding maintenance are: (1) How frequently will maintenance be required? (2) to what condition should a facility be allowed to deteriorate before any maintenance action is taken? and (3) What is the best maintenance alternative to take? Answers to these questions will vary for different facilities. While the methodology presented in this paper provides a framework for answering these questions for pavements, it also has potential applications to other types of deteriorating facilities.

The cost of maintaining a pavement depends directly on its condition or "state." Maintenance and rehabilitation funds are often allocated to pavements that are in the worst state or that exhibit an accelerated rate of deterioration. Factors that affect the rate of deterioration of pavements include traffic loads, weather, and construction techniques and materials. Observations of pavement condition vary considerably even when

[1]Asst. Prof., Dept. of General Engrg., Univ. of Illinois, 104 S. Mathews, Urbana, IL 61801.

[2]Assoc. Prof., Dept. of General Engrg., Univ. of Illinois, 104 S. Mathews, Urbana, IL 61801.

[3]Staff Engr., U.S. Army Corps of Engrs. Constr. Engrg. Res. Lab., 2902 Newmark Dr., Champaign, IL 61820.

[4]Production Control Mgr., General Motors Corp., Delco Products, Kettering, OH 45420.

[5]Engr., Abbott Lab., Abbott Park, IL 60045.

Note.—Discussion open until February 1, 1988. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on September 23, 1986. This paper is part of the *Journal of Transportation Engineering*, Vol. 113, No. 5, September, 1987. ©ASCE, ISSN 0733-947X/87/0005-0554/$01.00. Paper No. 21828.

such contributing factors are similar. Therefore, it is important that a procedure that can accommodate such randomness be incorporated into a pavement management system.

The objective of this investigation was to find a quantitative approach for finding the optimal sequence of maintenance expenditures that will satisfy certain performance objectives over a given planning horizon. The approach consists of three parts: (1) A mathematical model for pavement deterioration is specified in terms of a cumulative damage model based upon a Markov process; (2) the optimal maintenance decisions are found by probabilistic dynamic programming; and (3) a simulation is used to study the robustness and sensitivity of the optimal maintenance program. This integrated procedure can be used to consider different planning horizons, pavement types, and performance objectives. Further, as noted above, this same approach might be applied to other kinds of facilities, depending on their deterioration mechanisms.

The approach taken here permits the determination of the optimal maintenance decision for pavement in any state at any point in time in the planning horizon. During the sample calculation undertaken, it was generally found that the application of thin overlays when the pavement is only somewhat degraded is the least expensive approach. If too much deterioration is permitted before repair is initiated, the result is greater cost over the entire planning horizon. This finding resulted in the definition of a "minimum economical state" below which pavements should not be allowed to deteriorate. This state depends upon the deterioration model identified and the cost structure for the repair alternatives. An additional study was performed to determine the effect of delaying the optimal maintenance program, as might be necessary when the optimal budget exceeds the allocated funds. This approach permits the identification of the roads to which maintenance should be applied to reduce the added costs caused by delay. In addition, the total increase in cost caused by delay of optimal maintenance can be calculated for the planning horizon.

A discussion of the role of optimization in decision support for pavement management has been given by Liebman (1985). A number of approaches to pavement management, which vary considerably in sophistication, are described by Golabi, Kulkarni and Way (1982), Artman and Liebman (1983), Hoffman and Lyng (1985) and Novak and Kuo (1985). Some aspects of the methodology proposed in this paper were specifically directed towards the enhancement of PAVER, the pavement management system developed by Shahin and Kohn (1981) at the Construction Engineering Research Laboratory of the U.S. Army Corps of Engineers. For instance, the method used to identify the Markov process for pavement deterioration was based on the format of pavement condition data available in PAVER. Another approach might have been taken if the existing pavement management system and data base had been configured differently.

Elements of Methodological Approach

Markov Model for Pavement Deterioration

Methodology.—A probabilistic model for portraying pavement deterioration is required since the prediction of road conditions is a necessary part of deciding how to allocate maintenance funds. When data are available in the form of pavement condition versus age, approaches that have been taken include straight-line extrapolation, used by Shahin and Kohn (1981), and nonlinear regression, proposed by Hajek, et al. (1985). Extrapolation is essentially deterministic and does not attempt to explain what variability might be expected in future pavement conditions. Regression models are often chosen somewhat arbitrarily; polynomials and various exponential forms are manipulated to assume the general shape of the data. Little understanding of the process of deterioration is generated from such approaches. For instance, increasing variation in pavement condition with age, which is commonly observed, is not addressed by these "models."

A cumulative damage model based upon a Markov process has been proposed by Bogdanoff and Kozin (1985) and has been successfully used to investigate crack propagation and fatigue. The Markov process seems superior to the curve-fitting approaches mentioned above because it introduces a rational structure for interpreting road condition data. It can also be used to predict future pavement conditions in a probabilistic manner. The pavement management system developed by Golabi, Kulkarni and Way (1982) uses such a Markov process to describe the change in pavement conditions with time. Since their pavement data were available in the form of brief pavement histories, the identification of the Markov process took a different form from what is proposed here.

For the problem at hand, road condition data were available in the form of a survey in which many roads of the same type were inspected at a single point in time. The cumulative damage model proposed by Bogdanoff uses the concept of a duty cycle or repetitive period of operation during which damage can accumulate. In this study, pavement begins its life at some time in the past in near-perfect condition and then is subjected to a sequence of duty cycles that cause its condition to deteriorate. The duty cycle for pavement will be assumed to consist of one year's duration of weather and traffic load.

In an attempt to make the effect of the duty cycles on the pavement similar from year to year, pavements are categorized so that roads with similar pavement construction methods, traffic loads and geographical location are collected into one class or "family," as suggested by Nunez and Shahin (1986). Typical PCI data for a family of pavements tend to possess certain common features. In Fig. 1 it can be seen that, as the age of pavement increases, the average PCI worsens while the variation in PCI increases. These features of empirical data can be described by a Markov process with appropriately chosen parameters. In theory, as age increases, the variation in pavement conditions eventually decreases to zero, since all pavements approach the fully deteriorated state if they are not repaired.

The basic assumption for a Markov process is as follows: it is assumed that the decrement in PCI at the end of one duty cycle depends in a probabilistic manner only on the PCI at the start of the duty cycle and on that duty cycle itself, but is independent of how damage was accumulated up to the start of that duty cycle. This is a reasonable and unrestricting assumption. Furthermore, there are many methods for transforming processes that are apparently not Markov into Markov processes;

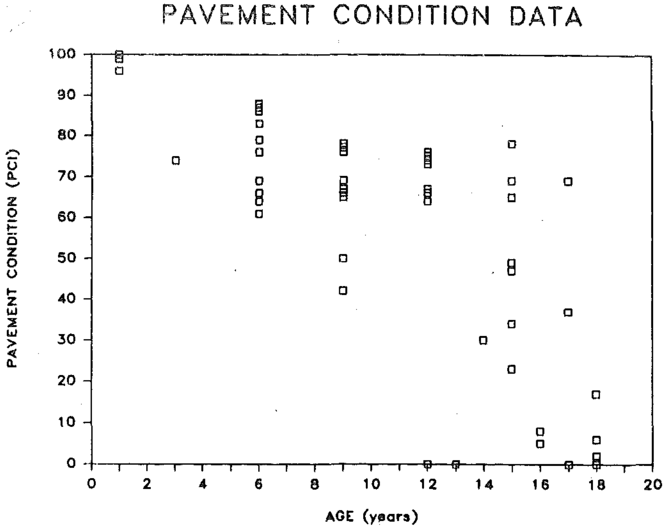

FIG. 1.—Pavement Condition Data Obtained from Pavement Survey

for example, the definition of the state vector can be modified as suggested by Cox and Miller (1965).

The pavement's state will be defined in terms of the PCI rating, which can vary from 0 to 100, where 100 denotes the best possible road condition. In this study, the states were defined as eight intervals of PCI, similar to those suggested in PAVER (see Table 1). This discretization of the state definition and the assumption of discrete time units in terms of duty cycles permits the deterioration process to be expressed as a Markov chain.

Since the determination of PCI in the field can be subject to an error of as much as ±5, the use of a much more refined state definition may be pointless. Past experience has shown that it is unlikely that the pavement condition will decrease more than 10 PCI in a single year. Thus, it is assumed that a pavement will either stay in its current state or jump to the next lowest state at the end of one duty cycle. Given that the

TABLE 1.—Pavement State Classification by PCI

PCI rating (1)	State representation (2)	State classification (3)
91–100	8	Excellent
81–90	7	.
71–80	6	.
61–70	5	.
51–60	4	.
41–50	3	.
21–40	2	.
0–20	1	Failed

pavement is in state j at the start of the duty cycle, let p_j be the probability of staying in state j and q_j be the complementary probability of going to the next state, $j - 1$, during a duty cycle. Mathematically, if $X(t)$ is the current state at time t, then

$$p_j = \text{prob}\,[X(t+1) = j \mid X)t) = j] \quad \text{(1a)}$$

$$q_j = \text{prob}\,[X(t+1) = j-1 \mid X(t) = j], \quad \text{(1b)}$$

$$\text{and} \quad q_j = 1 - p_j \quad \text{(1c)}$$

These probabilities can be arranged into a transition matrix, $\mathbf{P}$, which describes the probabilities of various possible state transitions:

$$\mathbf{P} = \begin{bmatrix} p_8 & q_8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_7 & q_7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_6 & q_6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_5 & q_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_4 & q_4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_3 & q_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & p_2 & q_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{(2)}$$

If the possibility of transition beyond the next lowest state during one duty cycle is significant, a third nonzero (upper) diagonal could be added to this matrix. This situation is more likely to occur if many more than eight states were used, corresponding to much smaller PCI intervals. In any event, each row of the transition matrix must sum to one. The entry "1" in the lower right hand corner corresponds to an absorbing state, from which pavement cannot escape without repair.

The state of damage at any time, t, is expressed in a probabilistic manner as a (1×8) row vector $\mathbf{p(t)}$, where

$$\mathbf{p}(t) = \{\text{prob}[X(t) = 8], \text{prob}[X(t) = 7], \ldots, \text{prob}[X(t) = 1]\} \quad \text{(3)}$$

and the elements of the row vector must sum to one. The initial state of the road is given by $p(0)$. For example, if it is known for certain that the pavement was in excellent condition when it was new, then

$$\mathbf{p}(0) = (1,0,0,0,0,0,0,0) \quad \text{(4)}$$

Other initial state vectors might be defined to reflect uncertainty in the initial condition of the road.

From Markov chain theory it is known that the state vector at some future time can be calculated from the transition matrix and the initial state vector, as

$$\mathbf{p}(1) = \mathbf{p}(0)\mathbf{P} \quad \text{(5a)}$$

$$\mathbf{p}(2) = \mathbf{p}(1)\mathbf{P} = \mathbf{p}(0)\mathbf{P}^2 \quad \text{(5b)}$$

$$\mathbf{p}(t) = \mathbf{p}(0)\mathbf{P}^t \quad \text{(5c)}$$

or, equivalently, as

$$\mathbf{p}(t+n) = \mathbf{p}(t)\mathbf{P}^n \quad \text{(5d)}$$

Although it may seem difficult to identify the transition matrix **P** from pavement data, the wealth of information contained in the state vector **p**(t) makes it a worthwhile effort. For example, the expected state and its standard deviation, along with probable bounds on pavement conditions, can be determined from **p**(t). The state transition mechanism required to implement the subsequent dynamic programming optimization can also be obtained from **P**.

Application.—Before proceeding with the identification of a transition matrix from pavement data, some observations will be made about the formulation. If a significant feature of the duty cycle were to change, it would be expected that a different **P** matrix would be employed to describe future deterioration. For example, if the traffic load changed from medium to heavy, a **P** matrix identified from a family of roads with heavy traffic loads, but with similar weather and construction type, would be used to model subsequent deterioration. If a pavement with its original surface received an overlay, subsequent deterioration would be modeled with a **P** matrix identified using data from a family of overlaid pavements with similar traffic loads and weather.

One important restriction that should not be overlooked is that the Markov chain is assumed here to be homogeneous. That is, the transition matrix **P** is assumed to be constant over time. Currently, the only attempt being made to account for nonhomogeneity is the allowing for the possibility of a change in the transition matrix after nonroutine maintenance. There is some concern about the assumption of homogeneity since traffic loads have gradually increased over time, providing a more destructive duty cycle in each subsequent year. Another factor that affects the validity of the homogeneity assumption is variation in maintenance policies and methods over time. It is recommended that the applicability of a nonhomogeneous Markov chain be the topic of future investigation.

To identify the transition matrix **P**, ideal data collection would consist of complete histories, or sample functions, of PCI versus time for roads belonging to a particular family of pavements. It would also be convenient if data were collected from roads put into use at the same point in time, so that the sequence of duty cycles would be identical for them. Pavement surveys do not supply data in this form since roads are placed into service at different times while the survey data are collected at a single point in time. The transformation into PCI versus age data is shown in Fig. 2. Thus, unfortunately, road survey data reflect different duty-cycle sequences and a substantial portion of the deterioration history is apparently not available from such data. Certainly, one byproduct of modeling deterioration in terms of a cumulative damage model is the identification of some of the shortcomings of available pavement data.

To find a transition matrix that reflects the available PCI versus age data, the following approach was taken. At each age for which data were collected, the difference between the expected state predicted from the Markov process and the observed pavement condition was calculated. Both the sum of the squared differences and the sum of the absolute value of the differences were used as objective functions in choosing the

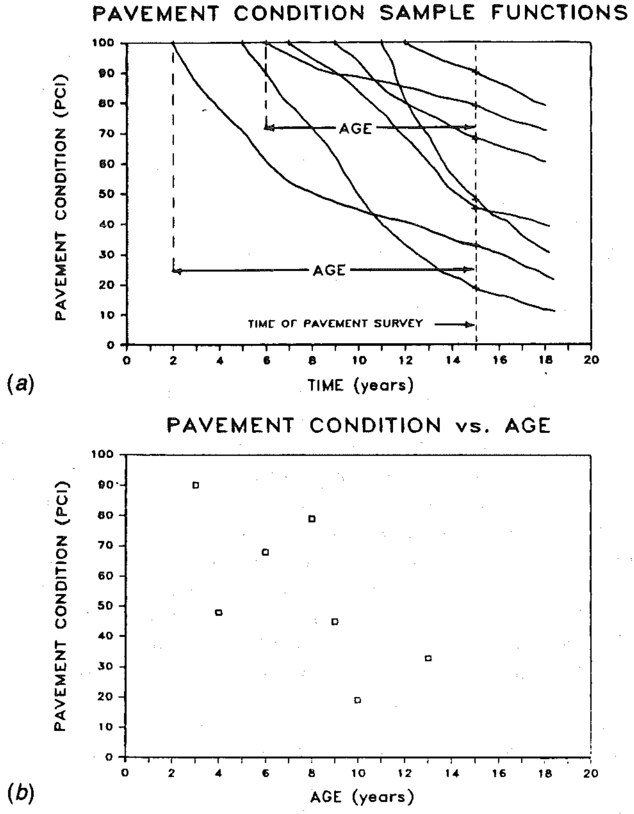

FIG. 2.—Transformation of Pavement Survey Data to PCI versus Age

elements of the transition matrix that produces the best fit. To accomplish the minimization of these objective functions, an algorithm developed by Fletcher and Powell (1963) was used.

Using two data sets from pavements at Great Lakes Naval Training Center, a number of interesting preliminary results were obtained. One data set represented original pavement constructed with asphalt concrete, while the other consisted of asphalt-overlaid pavement. Both bore secondary traffic loads. The results are given in Table 2, which contains the transition matrix diagonal elements. The trial number in the table refers to the results of using different initial guesses for the transition matrix elements. The estimates of the transition probabilities for the states of extreme deterioration appear to be somewhat sensitive to the initial guesses provided to the minimization algorithm. Currently there is an ongoing effort to find an identification procedure that produces more stable estimates. It is clear from the negligible changes in the objective function, however, that the expected pavement state is not greatly influenced by changes in p_2, p_3 and p_4. The elements p_5 through p_8 were found to have much greater influence on the expected state over the

TABLE 2.—Transition Matrix Diagonal Elements for Original and Overlaid Pavements

Trial number (1)	Objective function (2)	Transition Matrix Diagonal Elements						
		p_8 (3)	p_7 (4)	p_6 (5)	p_5 (6)	p_4 (7)	p_3 (8)	p_2 (9)
Original Pavement								
1	634.686	0.845	0.731	0.624	0.517	0.427	0.350	0.285
2	634.921	0.851	0.732	0.630	0.476	0.408	0.376	0.382
3	635.438	0.854	0.733	0.591	0.468	0.428	0.411	0.426
Overlaid Pavement								
1	802.762	0.844	0.719	0.507	0.303	0.303	0.202	0.292
2	805.633	0.808	0.704	0.583	0.443	0.431	0.437	0.463
3	803.231	0.823	0.710	0.605	0.503	0.301	0.201	0.103

range of pavement age represented by this data set.

The reason for this becomes clear upon examining the polynomial form of the state probabilities from which the expected states are calculated. For pavement starting in a new condition, the state probabilities at the end of one year depend on p_8 and p_7. At the end of the second year, the influence of p_6 is felt, and so on. In this way, the state transition probabilities for the deteriorated states have less influence on the expected state.

In Fig. 3 the observed pavement data are plotted with the expected value calculated from the transition matrix that minimized the objective

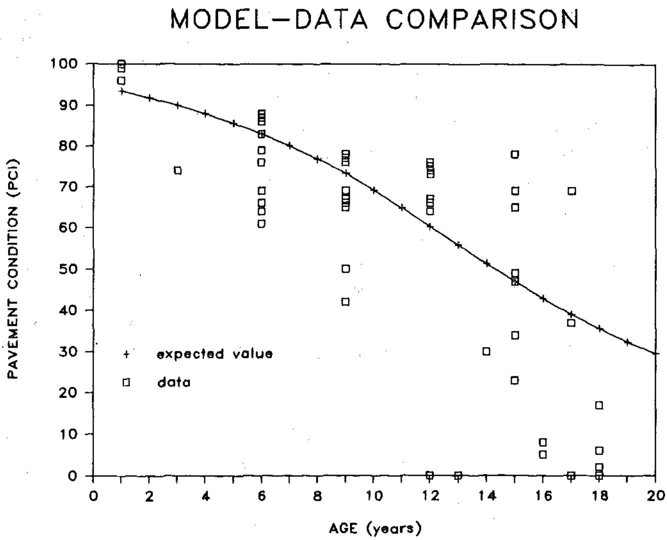

FIG. 3.—Comparison or Expected State from Cumulative Damage Model with Pavement Condition Data

TABLE 3.—Transition Matrices Diagonal Elements

Deterioration rate (1)	p_8 (2)	p_7 (3)	p_6 (4)	p_5 (5)	p_4 (6)	p_3 (7)	p_2 (8)
Original Pavement							
Better than average ($k = 1$)	0.882	0.707	0.684	0.626	0.860	0.882	0.891
Average ($k = 2$)	0.851	0.731	0.624	0.517	0.427	0.350	0.285
Worse than average ($k = 3$)	0.720	0.678	0.563	0.491	0.191	0.190	0.285
Overlaid Pavement							
Better than average ($k = 4$)	0.852	0.623	0.793	0.657	0.469	0.447	0.423
Average ($k = 5$)	0.844	0.719	0.507	0.303	0.303	0.202	0.292
Worse than average ($k = 6$)	0.711	0.651	0.551	0.415	0.230	0.014	0.001

function. Considerable variation from the expected state is observed for these data, even though the pavements are supposed to have been classified as belonging to a particular family. It is possible that this classification method had errors, e.g., perhaps the traffic loads on these pavements were not nearly equal, or perhaps pavements received different routine maintenance. Regardless, since the deterioration process appears stochastic, no classification scheme will account for all the variation observed in the data.

Because of this concern that the pavement data chosen for this study might have some classification errors, some additional transition matrices were identified from the data. Each data set was divided into two groups: above and below the mean or expected pavement state, as shown in Fig. 3. Additional transition matrices (see Table 3) were subsequently fit to each data set and used later in the investigation to describe pavements with different rates of deterioration. Better than average deterioration is represented by the transition matrix fit to the data above the mean pavement state predicted by the transition matrix identified using all the data. Worse than average deterioration is modeled by the transition matrix from data below the mean.

It should be emphasized that the expected value of the pavement state is calculated from the transition matrix only to obtain the best fit to pavement survey data. To predict the future condition of pavement, the transition matrix itself is used. For instance, if the overlaid pavement has a PCI of 75 in any particular year, the probability of staying in state 6 (PCI from 71 to 80) after one more year may be taken to be p_6 or 0.507. The state transition mechanism is much more obvious and accessible from the Markov chain than the regression "models" mentioned earlier.

Dynamic Programming Solution

Methodology.—Once the state transition can be modeled as a Markov process, the problem possesses a natural structure for which probabilistic dynamic programming can be used to find the optimal maintenance decisions. Using this approach, the expected cost of maintenance can be minimized over the entire planning horizon by performing a more tractable minimization associated with each year, beginning at the last year of the horizon. The recursive relationships necessary for the solution will

TABLE 4.—Set of Repair Alternatives

Repair alternative index, i (1)	Type of repair, R_i (2)
0	Routine maintenance
1	1-in. overlay
2	2-in. overlay
3	4-in. overlay
4	6-in. overlay
5	Reconstruction

be developed next. The transition matrices identified earlier and cost estimates for various repair alternatives were used to perform a sample computation of the optimal maintenance policy.

The example considers a planning horizon of T years where $t = 1, 2, \ldots, T$ denotes the year. Let $X^k(t) = j$, $(j = 1, 2, \ldots, 8)$ represent the state of the pavement in year t. The pavement classification k indicates that the governing transition matrix is $\mathbf{P}^k$, with diagonal elements denoted as $\{p_j^k\}$. For original (nonoverlaid) pavement, $k = 1, 2$, and 3 correspond to transition matrices fit to above mean data, all data, and below mean data, respectively. The pavement classifications $k = 4, 5$, and 6 correspond to the transition matrices fit using data from overlaid pavements. For example, $X^3(12) = 5$ refers to nonoverlaid pavement in state 5 in year 12 of the planning horizon that is deteriorating according to the transition matrix fit to pavement data below the mean. When pavement data do not exhibit as wide a variation as happened to be observed here, the pavement classification can be simplified; for example, the classifications $k = 2$ and $k = 5$ might be adequate. Reducing to this simple formulation is easily accomplished by means of the specification of a pavement classification transition probability, as will be shown later.

The repair alternatives considered for this demonstration are designated R_i, $i = 0, 1, \ldots, 5$ and are described in Table 4. For each repair alternative, R_i, there is a cost of implementation, $c[X^k(t), R_i]$, which depends on the current state of the pavement. The effects of inflation and interest can also make these costs functions of time. Both routine maintenance and pre-overlay costs become greater when pavements are in badly deteriorated states; thus the costs of all maintenance alternatives are a function of pavement state. The costs used in the sample calculation were taken from studies by Sharaf, et al. (1985) and Smith, et al. (1983) and are given in Table 5.

The pavement condition can be constrained in any year so that the pavement state $X^k(t)$ is kept above some minimum level $X_{\min}(t)$ with a specified probability, $p_{\min}(t)$. Of course, these constraints determine the feasibility of each maintenance alternative given the pavement state at each stage in the planning horizon. Such constraints may be enforced at every year of the planning horizon or, perhaps, at the last year only.

When a repair alternative is implemented, the possibility exists that a different transition matrix may govern subsequent pavement deterioration. The transition from pavement classification index k to k' will be expressed probabilistically. Let $q_{ik}^{k'}$ represent a pavement classification transition probability, where i represents the type of repair imple-

TABLE 5.—Costs of Repair Alternatives (in Dollars per Square Yard)

Pavement state (1)	Routine maintenance (2)	Overlays 1-in. (3)	2-in. (4)	4-in. (5)	6-in. (6)	Reconstruction (7)
8	0.04	1.90	3.81	7.61	11.42	25.97
7	0.15	2.00	3.91	7.71	11.52	25.97
6	0.31	2.20	4.11	7.91	11.72	25.97
5	0.65	4.73	6.64	10.43	14.25	25.97
4	0.83	7.15	9.06	12.86	16.67	25.97
3	1.40	8.78	10.69	14.49	18.30	25.97
2	2.00	10.40	12.31	16.11	19.92	25.97
1	6.90	19.90	21.81	25.61	29.42	25.97

mented, k represents the initial pavement classification, and k' represents the subsequent pavement classification after the repair R_i is completed. That is, the pavement classification transition probability is the probability that the transition matrix $\mathbf{P}^{k'}$ governs future pavement deterioration given that repair alternative R_i has been completed on pavement described by $\mathbf{P}^k$. These probabilities $q_{ik}^{k'}$ can be arranged into a pavement classification transition matrix $\mathbf{Q}_i$. As an example, noting that routine maintenance does not change the pavement classification, $\mathbf{Q}_0$ becomes the identity matrix.

The dynamic programming solution to this problem depends on the optimal decision in the last year of the planning horizon and on a recursive relationship that governs decisions in all previous years. Feasibility of the repair alternatives is decided in a straightforward manner. Nonroutine repair alternatives R_i $(i > 0)$ are always feasible since the pavement state is restored to $X^k(t) = 8$ when they are implemented. On the other hand, the routine maintenance alternative, R_0, must be examined to see whether either of the following conditions are satisfied to determine its feasibility:

$$X^k(t) - 1 \geq X_{\min}(t) \quad \text{(6a)}$$

$$\text{or} \quad X^k(t) = X_{\min}(t) \quad \text{and} \quad p_j^k \geq p_{\min}(t) \quad \text{(6b)}$$

After feasibility for routine maintenance has been determined, the minimum cost repair alternative is chosen, resulting in the optimal repair alternative $r^*[X^k(t)]$ with minimum cost $c^*[X^k(t)]$. Once the optimal repair is found for any year t for all $X^k(t)$, a recursive relationship can be developed to make optimal decisions for all subsequent stages. These stages correspond to previous years in the planning horizon since a "backwards pass" is being made with the dynamic programming algorithm, as explained in Keane and Wu (1985).

Let c_0 represent the expected cost of routine maintenance, while c' represents the minimum cost of all other repair alternatives. Of course, the repair alternatives must first be examined for feasibility using the probabilistic constraints mentioned earlier. If R_0 is infeasible, c_0 is set to infinity; otherwise, the mathematical expression for c_0 is

$$c_0 = q_{0k}^{k'}\{c[X^k(t), R_0] + p_j^k c^*[X^k(t+1)] + (1 - p_j^k)c^*[X^k(t+1) - 1]\} \quad \ldots\ldots (7)$$

The corresponding expression for c' reflects the probability of the pavement changing its pavement classification index (or transition matrix) as a result of a nonroutine repair alternative, R_i, being undertaken. Thus

$$c' = \min_{\substack{R_i \\ i>0}} \left[c[X^k(t), R_i] + \sum_{k'=1}^{6} q_{ik}^{k'}\{p_8^{k'} c^*[X^{k'}(t+1) = 8] \right.$$

$$\left. + (1 - p_8^{k'})c^*[X^k(t+1) = 7]\} \right] \quad \ldots\ldots (8)$$

Both c_0 and c' are expected costs over the remaining planning horizon, given pavement is found in state $X^k(t)$. The optimal repair cost for state $X^k(t)$ is computed as

$$c^*[X^k(t)] = \min(c_0, c') \quad \ldots\ldots (9)$$

while $r^*[X^k(t)]$ is subsequently defined as the minimum cost repair alternative. It is immediately clear that determining the optimal repair alternative in any year t requires the knowledge of the optimal repair alternative for each state in the following year. By starting in year T of the planning horizon, the optimal repair for years ($T - 1$, $T - 2$, ..., 2, 1) can be determined from this recursive approach.

Application.—A sample calculation was performed based on data provided by the U.S. Army Construction Engineering Research Laboratory; the details of the calculation are provided in Keane and Wu (1985). The state transition matrices $\mathbf{P}^k$, the costs for each repair alternative R_i, and the pavement classification transition matrices, $\mathbf{Q}_i$, must be specified. The state transition matrices and the repair costs are based on actual pavement data. In contrast, the $\mathbf{Q}_i$ were specified using expert judgment, using the principle that the rate of deterioration should remain the same or improve after repairs. Both the $\mathbf{Q}_i$ elements and the governing transition matrix after a particular repair could be identified in a more objective fashion, but more pavement deterioration data than is currently available would be required. An alternative approach would be to let $\mathbf{Q}_i$ reflect no change in deterioration rate after a maintenance action. However, this seems to be an unrealistically timid assumption.

The feasibility of repair alternatives available for each pavement state is also modified using expert judgment. For instance, for failed pavement, reconstruction is the only alternative designated as feasible. Overlays are not permitted since they will not restore the pavement to excellent condition, and using the overlaid pavement transition matrix subsequently would be incorrect. An example of a less clearcut judgment is that the 2-in. overlay was not allowed in this calculation unless pavement was at least in good condition. When pavement condition and maintenance history data become more extensive, more detailed identification of the various mechanisms of pavement transition can be undertaken. Then the dynamic programming algorithm can determine that particular actions are not optimal, rather than having them excluded exogenously from the feasible set. However, some expert judgment will always be required since data will not be forthcoming for certain impractical maintenance actions. Since failed pavements are not repaired with thin overlays, no data will become available and no transition matrix can be identified.

The repair costs (see Table 5) include material costs and pre-overlay operations and are found to be dependent upon the pavement state. The repair alternative representing reconstruction is assumed not to be state-dependent; its cost is the same for each pavement state $X^k(t)$. None of the costs are assumed to be dependent on the pavement classification index k, since the pavement classification represents the pavement deterioration rate rather than the pavement state.

The optimal maintenance program was obtained for a 20-yr planning horizon with a single constraint that the pavement state be at least 5 in year 20 with a probability of 0.95. The optimal maintenance actions and costs are given in Table 6 for pavements with an average rate of deterioration (k = 2 and 5). The repair alternative is given for each pavement state in each year of the planning horizon for both overlaid and non-overlaid pavement. It can be seen that there is a "minimum economical state" (MES) corresponding to $X^k(t) = 6$, below which pavement is not permitted to deteriorate. When the MES is reached, the thinnest overlay is applied and $X^k(t)$ is restored to 8. For pavements found below this state initially, various thicker overlays are applied to renew the pavements to state 8. For states above this minimum economical state, routine maintenance is performed. Once an overlay has been performed, the transition matrix for overlaid pavements governs; therefore, the cost will be the same for both original and overlaid pavements whose initial condition is less than the MES.

It is apparent that the optimal policy is unaffected by additional constraints in intermediate years that are below or equal to the MES. The constraint on the allowable pavement condition at the end of the planning horizon dictates whether or not an overlay is required in the last year.

A sensitivity analysis was performed on the structure of the costs used for this problem. Pre-overlay costs were made to vary as a multiple of routine maintenance costs; then the routine maintenance cost was made

TABLE 6.—Optimal Maintenance Action and Costs for 20-Yr Horizon for Pavement with Average Deterioration Rate (k = 2)

Pavement state (1)	R_i		Minimum Expected Cost ($/sq yd)	
	Years 1–19 (2)	Year 20 (3)	Original pavement (4)	Overlaid pavement (5)
8	0	0	3.32	3.40
7	0	0	4.94	4.97
6	1	0	5.66	5.66
5	2	2	10.06	10.06
4	3	3	16.00	16.00
3	3	3	17.69	17.69
2	5	5	28.58	28.58
1	5	5	28.58	28.58

a varying fraction of pre-overlay costs. The optimal repair strategy changed only when routine maintenance costs were much greater than pre-overlay costs, a case with no practical relevance. The characteristics of the optimal strategy appear to be generally in accordance with the findings of Darter, Smith and Shahin (1985): routine maintenance and thin overlays are preferred to major pavement rehabilitation with regard to total cost over the service life.

Simulation of Optimal Maintenance Program

Methodology.—After finding that the optimal policy was quite robust, the next step was to simulate the optimal policy and examine some of the variations that might be observed upon its implementation. Recall that dynamic programming guarantees only that the optimal maintenance policy will produce the minimum expected cost. No information is generated concerning cost variations that might be observed. Hence, dynamic programming gives little or no information concerning probable bounds that might be set on the optimal budget required during each year of the planning horizon.

There is a great deal of vital information that can be obtained by simulating the implementation of the maintenance policy that minimizes the expected cost over the planning horizon. The Markov process for deterioration can be easily simulated and the optimal maintenance policy applied to the resulting pavement state $X^k(t)$. Of course, if routine maintenance is chosen, the same transition matrix governs pavement deterioration during the next year or duty cycle. If an overlay or reconstruction is the optimal maintenance action, then the transition matrix governing further deterioration will change as dictated by the pavement classification index transition matrix, $\mathbf{Q}_i$. Since the simulation model is conceptually simple and easy to construct, the details are not described here and may be found in Keane and Wu (1985).

A limited number of studies were performed with the simulation model and certainly many more can be envisioned. The first effort was to obtain an estimate of the variation of the cost about its mean over the planning horizon. During this study, the frequency distribution for the timing of the required overlays was also obtained. The second study performed was the evaluation of the impact of delaying the optimal program, a situation that would arise when the optimal budget exceeds the allocated budget. This study proved to be particularly interesting since the practical ramifications seem to be significant. First of all, the optimal maintenance budget can be justified by quantifying the cost of deviating from it to satisfy short-term budget pressures. Secondly, the simulation results suggest a strategy for implementing the optimal budget by identifying the roads whose condition will cause the greatest cost increases over the planning horizon if not maintained optimally.

Application.—The first simulation study was performed and 1,000 replications were generated for each possible initial state. A replication refers to the generation of a sample sequence of pavement states, starting from a given initial state and continuing over the 20-yr planning horizon, according to the optimal repair policy and the appropriate governing Markov chains. The sample mean, $\bar{c}$, and standard deviation for cost, s, as they vary with the initial state are given in Table 7. The mean

TABLE 7.—Simulated Costs for 20-Yr Horizon (in Dollars per Square Yard)

Pavement state, j (1)	Dynamic programming expected cost, c^* (2)	Simulation mean cost, $\bar{c}$ (3)	Simulation standard deviation, s (4)
Original Pavement with Average Deterioration Rate (k = 2)			
8	3.32	3.24	1.56
7	4.94	4.95	1.49
6	5.66	5.63	1.58
5	10.06	10.10	1.61
4	16.00	16.02	1.67
3	17.69	17.22	1.65
2	28.58	28.74	1.71
1	28.58	28.72	1.66
Overlaid Pavement with Average Deterioration Rate (k = 5)			
8	3.40	3.46	1.59
7	4.97	5.04	1.49
6	5.66	5.62	1.60
5	10.06	10.09	1.66
4	16.00	16.05	1.57
3	17.69	17.71	1.62
2	28.58	28.63	1.69
1	28.58	28.61	1.66

costs are in very close agreement with the expected costs determined by dynamic programming since a 95% confidence interval is $\bar{c} \pm 1.96s\sqrt{1,000}$. As expected, the mean cost became greater as the initial condition of the road became worse. What may not have been expected is that the standard deviation is roughly independent of the initial road condition. Some consideration of the nature of the optimal policy can explain this finding. Pavements found below the minimum economical state are immediately repaired and thus placed in state $X^k(t) = 8$. Subsequently, they are repaired in the same manner as new pavement to prevent them from falling below the minimum economical state. Thus, although the mean costs differ greatly depending on the initial state, the variation about this mean cost is of approximately the same magnitude.

By combining this information with a road survey or inventory of pavement conditions, an expected budget can be estimated for an existing network of roads. By making a distributional assumption (for example, that costs are normally distributed), some confidence intervals can be placed on the optimal budget required. Since the required budget in each year is the sum of the optimal budget for each type of pavement weighted by the number in various states of pavement condition, the assumption of normality for the total budget may be plausible. If the normality assumption is not palatable, an empirical estimate of the distribution of costs in each year can be obtained from the simulation. In this same fashion, various probabilistic estimates for the required optimal maintenance budget can be calculated without employing any distributional assumptions.

A related statistic also of interest is the frequency at which overlays are required in each year of the planning horizon for each pavement initial state. The number of 1-in. overlays and the years in which they occurred were recorded during 1,000 replications of the simulation. An example histogram for original pavement with average deterioration ($k = 2$), initially in excellent condition ($j = 8$) is given in Fig. 4. After five years the overlay frequency becomes quite uniform, implying that anticipated budgets would be uniform, also. This same feature is observed for other deterioration rates and initial conditions, indicating that the optimal maintenance budgets may be uniform after the first few years elapse.

The second study performed was the evaluation of the impact of delaying the optimal program. The simulation was run to find the economic consequences over the planning horizon of delaying the optimal policy by 1, 2, and 3 years. During the years of delay, routine maintenance was assumed to be performed. This assumption requires that the allocated budget is at least adequate for the performance of the routine maintenance alternative. It is also assumed that after the delay, the allocated budget is sufficient to resume the optimal maintenance program. The calculation of the expected cost of delay could have been accomplished using the information generated during the dynamic programming solution. Instead, simulation was used to obtain an estimate of the variation in cost in addition to its expected value.

The results of the simulations are given in Fig. 5 for various initial pavement states. It can be seen that the percentage increase in cost is less for pavements in good condition since the deviation from the optimal program is slight. For pavements near the minimum economical state the percentage increase in the mean cost is quite large, as is the

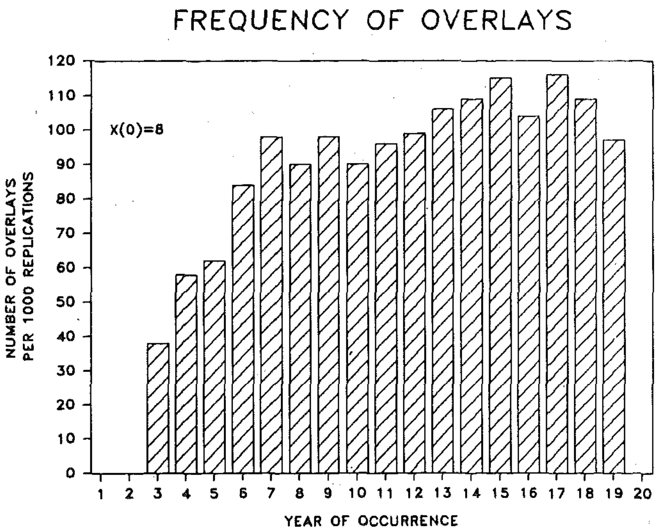

FIG. 4.—Histogram for Year of Occurrence of Overlay during 1,000 Replications of Simulation

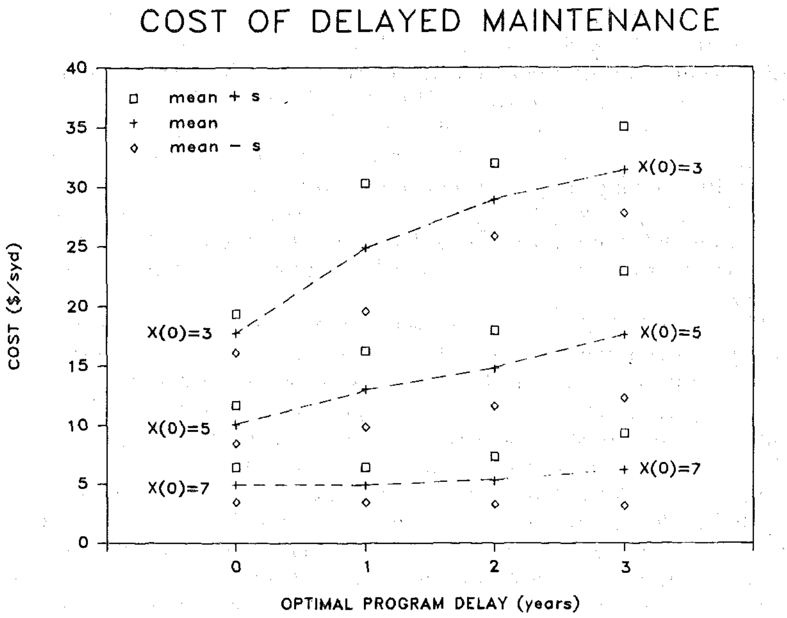

FIG. 5.—Cost of Maintenance of 20-yr Horizon after Delay of Optimal Program

increase in the standard deviation. The delay increases the probability of needing a thicker overlay than would have been required under the optimal maintenance program. Because deterioration is not deterministic, the requirement for this thick overlay is not certain; thus the variation of cost increases.

For pavements initially in highly deteriorated states (e.g., $X^k(0) = 3$), the increase in mean cost is rapid for a one-year delay and then slows down, with similar behavior being seen in the variation in cost. For this particular case, great uncertainty lies in whether the road will require a thick overlay or complete reconstruction after the delay. After a delay of one year (during which expensive routine maintenance must be is performed), reconstruction will be required for a significant fraction of the pavements during the next year. Therefore, both the mean cost and its variation increase. After another year of delay the specter of reconstruction becomes a reality for many of the pavements and the variation about this increased mean cost becomes less as this outcome becomes more certain.

There are a number of interesting results contained in the statistics generated by the simulations of the delay in the optimal maintenance program. Some of them are anomalies associated with the particular feasible set of repair alternatives and must be examined in more detail. A few other peculiar results are related to the $\mathbf{Q}_i$ matrices, which indicate changes in the transition matrix after a repair alternative is implemented. The results discussed above are more general and demonstrate the useful information that a decision-maker can generate with a simulation model.

Concluding Analysis

The approach proposed here to determine the optimal maintenance program for a deteriorating facility turned out to be successful when applied to pavement maintenance. A Markov chain can be identified as an appropriate cumulative damage model that exhibits the salient features of pavement deterioration data. Dynamic programming provides a straightforward and efficient algorithm for finding the minimum cost repair policy given a planning horizon and constraints on pavement condition. Simulation was used to study model sensitivity and collect vital statistics for decision-making when the optimal policy is applied to sample road condition histories.

Some interesting results were generated during the sample computation. A "minimum economical state" was found, which characterized the optimal repair policy found by dynamic programming. It was also demonstrated that such an integrated approach can also be used to quantify the costs of delaying the optimal program and thus justify a maintenance budget.

The use of an index (like PCI) to define pavement state causes some concern because it is an aggregation of specific distress measurements. Maintenance requirements and therefore costs are more directly related to these components of the index. It is also possible that future pavement conditions might be more accurately predicted from the components of the index. As one of the referees pointed out, there could be an enormous amount of work involved with identifying Markov chains from multiple distresses rather than a simple index. Certainly this should be an area of further investigation; currently there is an ongoing effort to explore the relationship between maintenance cost and PCI.

Although there seems to be a potential for further enhancement, the methodology presented here demonstrates a considerable potential for estimating budget requirements, particularly for a network of pavements. The approach can be implemented with a pavement data base to immediately create a pavement management system with considerable utility.

Acknowledgment

The writers would like to express their appreciation to the U.S. Army Corps of Engineers Construction Engineering Research Laboratory, which provided financial support so that this study could be undertaken.

Appendix.—References

Artman, D., and Liebman, J. (1983). "Optimization of long-range major rehabilitation of airfield pavements." *Transp. Res. Record 938*. Jan., 1–11.

Bogdanoff, J., and Kozin, F. (1985). *Probabilistic Models of Cumulative Damage*. John Wiley and Sons, Inc., New York, N.Y.

Cox, D., and Miller, H. (1965). *The Theory of Stochastic Processes*. John Wiley and Sons, Inc., New York, N.Y.

Darter, M., Smith, R., and Shahin, M. (1985). "Use of life cycle cost analysis as

the basis for determining cost-effectiveness of maintenance and rehabilitation treatments for developing a network level assignment procedure." *Proc. North American Pavement Mgmt. Conf.* 2, Toronto, Ontario, Mar., 7.10–7.18.
Fletcher, R., and Powell, M. (1963). "A rapidly convergent descent method for minimization." *Computer J.*, 6, 163–168.
Golabi, K., Kulkarni, R., and Way, G. (1982). "A statewide pavement management system." *Interfaces* 12(6), 5–21.
Hajek, J., Phang, W., Prakash, A., and Wrong, G. (1985). "Performance prediction for pavement management." *Proc. North American Pavement Mgmt. Conf.*, 1, Toronto, Ontario, Mar., 4.123–4.134.
Hoffman, G., and Lyng A. (1985). "Pennsylvania's planning process using STAMPP needs information." *Proc. North American Pavement Mgmt. Conf.*, 1, Toronto, Ontario, Mar., 3.14–31.
Keane, P., and Wu, M. (1985). "An integrated decision-making methodology for optimal maintenance strategies." *Dept. of General Engrg. Tech. Report*, Univ. of Illinois, Urbana, Ill.
Liebman, J. (1985). "Optimization tools for pavement management." *Proc. North American Pavement Mgmt. Conf.* 2, Toronto, Ontario, Mar., 6.6–6.15.
Novak, E., and Kuo, W. (1985). "Michigan department of transportation's pavement management system." *Proc. North American Pavement Mgmt. Conf.* 1, Toronto, Ontario, Mar., 4.97–4.105.
Nunez, M., and Shahin, M. (1986). "Pavement condition data analysis and modeling." Presented at Transp. Board Annual Meeting, Washington, D.C., Jan.
Shahin, M., and Kohn, S. (1981). "Pavement maintenance management for roads and parking lots." *Tech. Report M-294*. Constr. Engrg. Res. Lab., U.S. Army Corps of Engrs., Champaign, Ill.
Shahin, M., and Kohn, S. (1982). "Overview of the PAVER pavement management system." *Tech. Report M-310*. Constr. Engrg. Res. Lab., U.S. Army Corps of Engrs., Champaign, Ill.
Sharaf, E., Reichelt, E., and White, T. (1985). "Development of a cost model for the PAVER system." *USA-CERL Proj. DAC W88-83-D-0003, Progress Rep.*, School of Civ. Engrg., Purdue University, W. Lafayette, Ind.
Smith, R., Darter, M., Zimmer, T., and Carpenter, S. (1983). "Architectural-engineering services for pavement inspection and evaluation of the Great Lakes Naval Training Center, Illinois." *Final Report, N62472-82-C-6112*. ERES, Inc., Champaign, Ill.

Appendix II.—Notation

The following symbols are used in this paper:

$c[X^k(t), R_i]$ = cost of repair i when pavement is in state $X^k(t)$;
$\bar{c}$ = sample mean cost generated by simulation;
j = index for pavement state ($j = 1, \ldots, 8$);
k = index for pavement classification ($k = 1, \ldots, 6$);
$\mathbf{P}^k$ = pavement transition matrix **k**;
$\mathbf{p}(t)$ = probabilistic state vector for pavement condition;
p_j^k = diagonal element of transition matrix **k**;
$q_{ik}^{k'}$ = pavement classification transition probability;
R_i = repair alternative i, ($i = 0, \ldots, 5$);
$r^*[X^k(t)]$ = optimal repair alternative when pavement is in state $X^k(t)$;
s = sample standard deviation for cost generated by simulation;
t = year of planning horizon; and
$X^k(t)$ = pavement condition in year t.