Economy. In Highway Research Record 77, HRB, National Research Council, Washington, D.C., 1965, pp 1-18.
12. P.J. Claffey. Running Costs of Motor Vehicles as Affected by Road Design and Traffic. NCHRP Report III. HRB, National Research Council, Washington, D.C., 1971, 97 pp.
13. J. De Weille. Quantification of Road User Savings. World Bank Staff Occasional Paper 2. World Bank, Washington, D.C., 1966.
14. J. Zaniewski et al. Vehicle Operating Costs, Fuel Consumption and Pavement Type and Condition Factors. Report FHWA-RD-80. Texas Research and Development Company, Austin, 1981.
15. B.C. Butler, Jr. Use Related Vehicle Depreciation. In Transportation Research Record 955. TRB, National Research Council, Washington, D.C., 1984, pp. 1-9.

Publication of this paper sponsored by Committee on Maintenance and Operations Management.

Optimal Rehabilitation Frequencies for Highway Pavements

MICHAEL J. MARKOW and WAYNE S. BALTA

ABSTRACT

The maintenance and rehabilitation of existing, mature facilities are becoming increasingly important components of highway activity. Yet, although the planning, budgeting, evaluation, and management of maintenance and rehabilitation are different from corresponding actions for new construction, comparatively little work has been devoted to the development of planning and management tools intended specifically for repair programs. For a number of reasons, the optimization of maintenance and rehabilitation policy is difficult, and new concepts and analytic approaches need to be formulated to address this problem. Recently, the usefulness of dynamic control theory for optimizing transport investment decisions has been demonstrated. Control theory structures a problem in terms of a dynamic (i.e., time varying) objective function (e.g., maximize total transport-related benefits over time) subject to dynamic constraints (e.g., equations describing changes in pavement condition due to deterioration and repair or variations in traffic levels responding to current pavement condition). The several factors that influence the problem are structured in terms of state variables (over which decision makers have no control, such as traffic, weather, and soil) and control variables (over which decision makers exercise judgment, such as maintenance and rehabilitation policy). Dynamic control theory thus presents an attractive analytical tool for management of highway infrastructure; it encompasses all the key variables of interest, allows technically correct engineering and economic relationships to be expressed in problem formulation, and leads directly and efficiently to solution of optimal maintenance and rehabilitation policy. The tenets of dynamic control theory are described, and a numerical example of the use of dynamic control theory to optimize the overlay frequency on highway pavements in the United States is given.

The maintenance and rehabilitation of existing, mature facilities are becoming increasingly important components of highway activity. However, comparatively little work has been devoted to the development of planning and management tools intended specifically for repair programs. Yet, decisions regarding the planning, budgeting, evaluation, and management of maintenance and rehabilitation are different from corresponding actions for new construction:

1. Planning and managing maintenance and rehabilitation programs require an understanding of concepts underlying facility performance, as opposed to facility design.
2. There is a need to understand the role of

maintenance and rehabilitation in influencing facility performance.

3. The planning of maintenance programs implies an ability to evaluate life-cycle performance and costs, with trade-offs measured in economic as well as technical terms.

4. Decisions to repair existing facilities are complicated by the wide range of activities possible (ranging from minor routine maintenance to major rehabilitation or reconstruction), problems in spatial and temporal allocation of resources throughout a network, and choices between investment and noninvestment policies (e.g., pavement strengthening versus adjustments in size and weight limits).

5. For those facilities that do not fail catastrophically (e.g., highway pavements, rail track), it is difficult to define the point of failure, which, in turn, complicates the specification of standards governing facility performance, safety, and cost.

As a result of these characteristics, the optimization of maintenance and rehabilitation policy is difficult, and new concepts and analytic approaches need to be introduced to address this problem. A first step in this direction was taken by Fernandez (1), who demonstrated the applicability of dynamic control theory to transportation investments. Fernandez used this approach to solve for several general case studies in transportation, including stage construction and routine maintenance, but not rehabilitation.

From an engineering point of view, rehabilitation is a major concern of highway administrators and a key element of pavement 3R programs across the nation. From an analytic point of view, however, rehabilitation poses the problem of discontinuities in the pavement condition history, as shown in Figure 1. These discontinuities arise through cycles of deterioration and major repair or renewal and signal the expenditure of significant sums of money. These discontinuities complicate the representation of facility performance and costs over time, as well as the determination of the optimal time to rehabilitate.

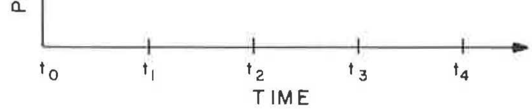

FIGURE 1 Effect of rehabilitation on pavement deterioration.

Recently, the necessary assumptions and mathematical solutions needed to address rehabilitation were successfully formulated by Balta (2). Balta developed his model, based on principles of control theory for discontinuities in state variables and system equations at interior points, to solve for the optimal timing of rehabilitation of a highway pavement. It is this solution, applied to pavement overlays, which is described in this paper.

The control theory approach enjoys several distinct advantages:

- It provides a unified conceptual and methodological framework for addressing rehabilitation policy;
- It features a closed-form optimization procedure;
- It encompasses all of the relevant variables pertaining to the demand-responsive approach to maintenance, including an interaction between demand for use and quality of pavement; and
- Although it has been developed as a policy model based in part on elements of economic and utility theory, it is intuitively appealing from an engineering point of view.

Although the mathematics of control theory may appear intimidating, the derived solution is elegant and leads to a surprisingly practical set of curves for optimal rehabilitation intervals that can be easily applied in the office or in the field. The formulation of the control theory model will be described first. Then, application of the model to the optimal timing of rehabilitation, using pavement overlays as examples, will be illustrated.

MODEL FORMULATION

Problem Description

The mathematical framework of optimal control theory can be used to model any number of stages between discontinuities for any length of time desired. For simplicity, however, the rehabilitation investment decision model developed in this paper will address only two such stages: before and after an investment is made. Given that the investment occurs at time t^*, Stage 1 then occurs between the beginning of the analysis period (time t_0) and the moment just before the investment (time t^{*-}). Similarly, Stage 2 occurs between the moment just after the investment (time t^{*+}) and the end of the planning horizon (time T). It is nonetheless clear that the model can be generalized for any number of investments. Figure 2 shows a summary of the pavement behavior that is analyzed in this model.

Like other optimization procedures, an optimal control model is specified with system variables, an objective function, and constraints. The state variables describing the system are defined by ordinary differential equations; these equations describe the state of the system at any time t. In this paper the system described is a highway pavement. The variables that characterize the system are

- $S(t) \equiv$ quality or condition of the pavement and
- $q(t) \equiv$ demand or traffic volume on the pavement.

Pavement condition, or S(t), is measured in the model in terms of the present serviceability index (PSI) as defined by AASHTO (3). Traffic demand, or q(t), is measured in the model in terms of 18-kip equivalent single axle loads.

Pavement Condition

Mathematically, pavement condition, S(t), is continuous during each stage between investments but it is discontinuous at the time of the investment. Conse-

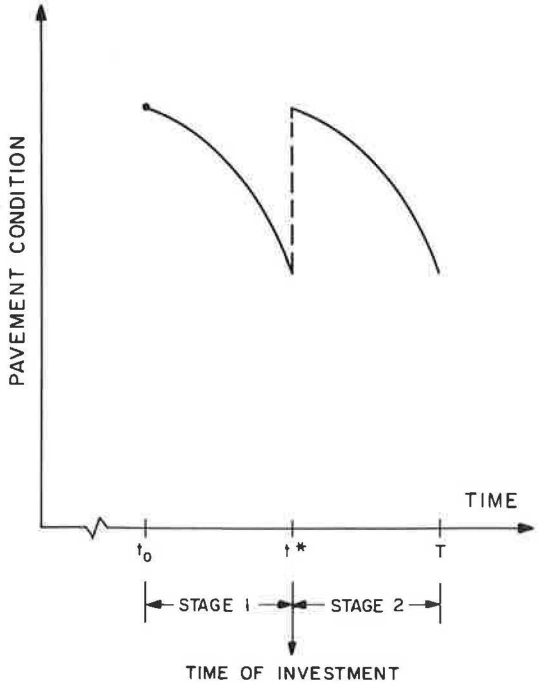

FIGURE 2 Pavement behavior analyzed in rehabilitation model.

quently the change in pavement condition over time, $\dot{S}(t)$ (where the dot denotes a time derivative), is also continuous during each stage between investments but discontinuous at the time of the investment. For the rehabilitation investment decision model it is assumed that $S(t)$ decays from its initial value in a nonlinear fashion as shown in Figure 3. $\dot{S}(t)$ will behave similarly. (Note that either concave or convex relationships may be used to represent different mechanisms of pavement damage.) It is further assumed that $S(t)$ and $\dot{S}(t)$ depend explicitly on time. Appropriate equations that capture this behavior are

$$\dot{S}_1(t) = -\beta_1 k_1 e^{\beta_1 t} \qquad t\epsilon(0,t^{*-}) \tag{1}$$

and

$$\dot{S}_2(t) = -\beta_2 k_2 e^{\beta_2 (t-t^*)} \qquad t\epsilon(t^{*+},T) \tag{2}$$

where the subscripts 1 and 2 refer to stages one and two, respectively, and t = 0 corresponds to the beginning of the analysis period t_0. $S(t)$ is equal to the integral of $\dot{S}(t)$ and can be expressed as

$$S_1(t) = k_1 + S_0 - k_1 e^{\beta_1 t} \qquad t\epsilon(0,t^{*-}) \tag{3}$$

and

$$S_2(t) = k_2 + S_A - k_2 e^{\beta_2 (t-t^*)} \qquad t\epsilon(t^{*+},T) \tag{4}$$

In these equations, β_1 and β_2 are parameters that control how quickly pavement condition deteriorates; as β becomes larger, the condition decays at a faster pace. Values for β depend on several factors including pavement design, quality of construction, traffic volume and composition, weather, soil and drainage conditions, and extent of routine maintenance performed. It can be argued that $S(t)$ should depend explicitly on each of these factors. For mathematical simplicity in the development of this model, however, all of these factors are assumed to be implicit in a given value of β. Under these circumstances, these factors should be accounted for when formulating pavement damage as a time-dependent function. The benefit of mathematical simplicity will become evident as the model unfolds. Increased analytical complexity is really the only price to be paid for a more explicit approach.

The terms k_1, k_2, S_0, and S_A in Equations 1-4 are constants. S_0 and S_A represent the initial level of pavement condition and the level of pavement condition immediately following rehabilitation, respectively. (It is assumed that $S_A \leq S_0$.) The constants k_1 and k_2 interact with β_1 and β_2 to further affect the extent to which the deterioration function deviates from a linear equation. Note that if t = 0 in Equation 3, $S_1(0) = S_0$. If $t = t^*$ in Equation 4, $S_2(t^*) = S_A$. Because they are all constants, k_1 and S_0 as well as k_2 and S_A could be combined into one value. The reason for expressing them separately is to preserve the identity of S_0 and S_A along with the interaction between β_1 and k_1 or β_2 and k_2.

This discrimination permits straightforward sensitivity analyses. For example, it may be thought that a pavement deteriorates very little over t years and then suddenly decays very quickly. Alternatively, it may be thought that a pavement deteriorates at a more constant pace over the same time period. The functional form for $S(t)$ in this model is easily adaptable to different combinations of k and β.

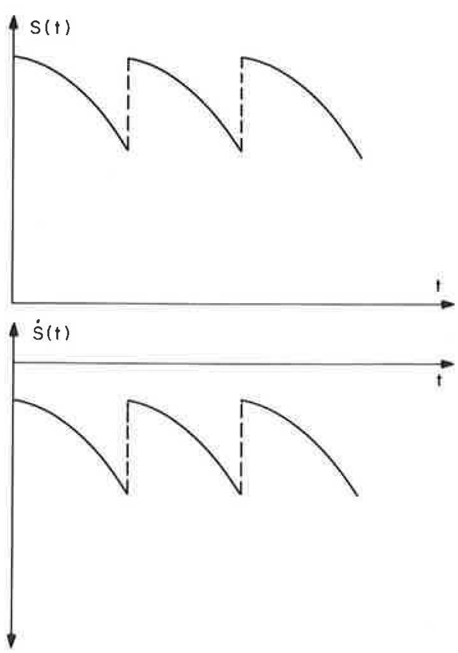

FIGURE 3 Pavement deterioration functions for rehabilitation model.

Traffic

The other variable characterizing the system of a highway pavement is $q(t)$, traffic demand. Fernandez (1) asserted that it is important to account for an interaction between the quality of a pavement and the demand for its use. Mathematically this relationship is represented by a differential equation of the form:

$$\dot{q}(t) = a(t)\,S(t) + b(t) \tag{5}$$

meaning that demand changes over time due to the condition of the pavement as well as to some external growth factor. In the first term, a(t) is a function that indicates how the number of road users changes due to a unit change in pavement quality. The function b(t) represents that part of the traffic growth rate that is independent of pavement quality. For mathematical simplicity, once again, the functions a(t) and b(t) are assumed to be constants in the development of this model.

Objective Function

The next step is to establish the objective function. The objective of the rehabilitation investment decision model is to maximize the present value of the net benefits derived from operation of the road over some planning period. Mathematically this is stated as follows:

$$\text{Maximize } J = \int_0^{t^{*-}} [U(t) - C(S,q)]\; q(t) \exp(-\rho t)$$

$$- I[S(t^{*-}), S(t^{*+}), t^*] \exp(-\rho t^*)$$

$$+ \int_{t^{*+}}^{T} [U(t) - C(S,q)]\; q(t) \exp(-\rho t) + \psi\,[q(T), T] \tag{6}$$

U(t) is the benefit that each user perceives from using the road at time t. C(S,q) is the cost that each user experiences from using the road. It is assumed that user costs depend on the quality of the pavement and the number of road users. Because user utility and user cost are specified on a per user basis, multiplying their difference by the total number of users q(t) yields the total private net benefit obtained at any time. The term $\exp(-\rho t)$ is a discount factor that converts all values to present value; ρ is the appropriate discount rate. Integrating from 0 to t^{*-} represents a summation of the net benefits during Stage 1.

The term $I[S(t^{*-}), S(t^{*+}), t^*]$ represents the cost of the investment. The magnitude of this investment may be a function of many different factors. Here it is shown as a function of the pavement condition just before the investment, the desired new pavement condition following the investment, and the time of investment itself. Having $I \sim f[S(t^{*+})]$ is particularly relevant for the case of reconstruction where the investment would depend on the desired reconstructed quality of the pavement. Having $I \sim f(t^*)$ makes it possible to account for any real price changes in the cost of rehabilitation. For the case of overlays, it is reasonable and sufficient to use $I \sim f[S(t^{*-})]$. Existing condition affects overlay cost at least in accordance with required pavement preparation, which depends on the extent of crack sealing, patching, and localized pavement repair that is needed before the new wearing course is placed. Furthermore, because pavement condition is represented by AASHTO's present serviceability index in this model (where the rehabilitation is an overlay), the surface condition will be essentially the same following most overlays regardless of cost. The model is therefore explicitly developed using $I[S(t^{*-})]$ as the functional form for investment cost.

The third term of the objective function, $\int_{t^{*+}}^{T}$ $[U(t) - c(S,q)]\; q(t) \exp(-\rho t)$, is identical to the first except that it represents a summation of the net benefits during Stage 2. The final term, $\psi[q(T),T]$, represents the salvage value of the pavement at the end of its useful life. This form implies that salvage value of the pavement depends in some manner on the number of users at the terminal time T. Later in this paper it is specifically assumed that ψ depends on an infinite stream of user benefits beginning at T. This assumption is largely for convenience. Highway planners may justly assert that salvage value should depend on terminal pavement condition. The only consequence of incorporating this consideration into the model would be increased mathematical complexity.

There are two additional points to make about the objective function. First, it is predicated on the concept of maximizing net benefits. (The model could also be reformulated with the objective of minimizing net costs.) Second, the objective function rests on microeconomic principles of maximizing the net social benefits associated with a public facility. This is highlighted by the inclusion of user costs, which play a significant role in shaping the model's investment policy.

Solution

The problem to be solved may be summarized as follows: to determine the optimal time of rehabilitation, t*, that maximizes the objective function, Equation 6, subject to the technical constraints represented by Equations 1, 2, and 5. The solution to this problem involves computing the Hamiltonian function. The relationship of the Hamiltonian to dynamic control models is analogous to the relationship of the Lagrangian to static optimization models and involves the specification of adjoint variables analogous to Lagrangian multipliers. The details of the solution become quite involved, and are explained fully elsewhere (2). Moreover, Balta has developed individual solutions for different assumptions of traffic and investment cost (i.e., whether these are constant or variable over time). The focus in this paper will be on the most general solution obtained, considering variable traffic (i.e., allowing for nonzero rates of traffic growth and decline) and variable investment cost (i.e., implying a relationship between overlay thickness or cost and time-varying pavement condition).

The solution is expressed in the form of a decision rule, evaluated at t*, balancing marginal costs and marginal benefits. To obtain an explicit rule, however, it is necessary to specify relationships governing pavement deterioration over time, user costs (as functions of both traffic volume and pavement condition), traffic growth over time, rehabilitation cost over time (as a function of pavement condition when overlaid), and salvage value. Some of these relationships have already been defined in the problem description: pavement deterioration is characterized by Equations 1 and 2 or 3 and 4, and traffic growth is governed by Equation 5. Following are explanations of the remaining relationships used in the solution.

User Costs

This model considers two components of user costs:

- Vehicle operating costs and
- Travel time costs.

These components are both based on the research reported in the EAROMAR-2 simulation program (4). Vehicle operating costs depend on fuel, oil, and tire consumption. Travel time costs depend on the user's trip purpose and income level.

The initial form of the user cost function tested in the model was

$$C[S(t),q(t)] = c - \epsilon S(t) + \delta q(t)^n \quad (7)$$

The first term of the user cost function, c, represents the user cost associated with low levels of traffic volume at a pavement condition of zero. The second term, $-\epsilon S(t)$, shows user cost decreasing linearly by ϵ for a unit increase in pavement quality. The third term, $\delta q(t)^n$, implies that the rise in user cost due to increased traffic (congestion) is a power function.

After applying this model in some test runs, however, it was noted that some of the results were unrealistic in light of normal practices of transportation and highway departments: rehabilitation was occurring too early in the pavement's life, after only a small decline in S(t). It was concluded that the most likely element causing these results was the assumed linear relationship between user costs and pavement surface condition and the magnitude of its slope, as shown in Figure 4. These values were determined using successive EAROMAR-2 simulations. What makes this relationship unrealistic is the magnitude of the increased cost experienced for a unit decline in PSI within the range of PSIs normally encountered on pavements in service (from approximately 4.5 to 2.0 PSI). As a result, the potential savings in user costs is so great that this marginal benefit will equal the marginal cost of undertaking the investment at a very slight degradation in surface condition.

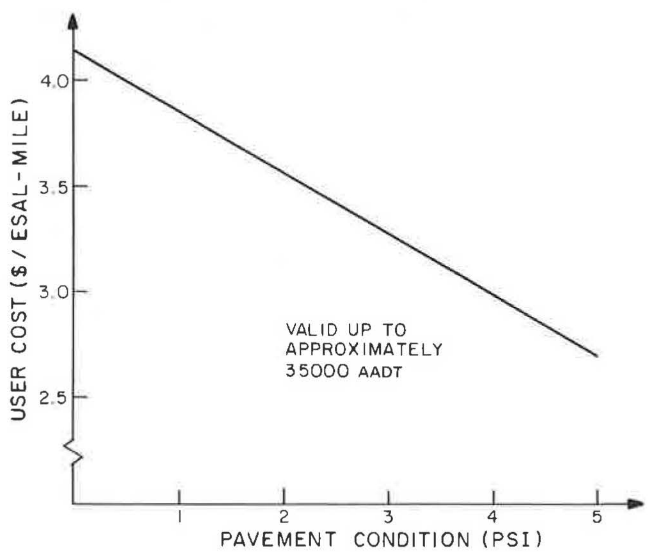

FIGURE 4 Linear C(S) relationship used to obtain initial results.

However, recent research demonstrates that the change in user cost (specifically in vehicle operating cost) experienced due to a unit change in PSI within the range of PSIs normally encountered on pavements in the United States is not much at all and is, indeed, significantly lower than the values shown in Figure 4 (5,6). Moreover, Ross (6) asserts that the true relationship between user costs and pavement surface condition is nonlinear.

The optimal rehabilitation investment decision model therefore employs a revised, nonlinear C versus S relationship as shown in Figure 5. This curve is expressed by the following natural antilogarithmic function:

$$C(S) = \mu C_0 e^{-\omega S(t)} + (1-\mu) C_o \quad (8)$$

where

C_0 = user cost for travel on the worst possible pavement ($S \to 0_{PSI}$) without traffic congestion and
ω,μ = parameters controlling the shape of the function.

The parameter μ plays a role similar to that of k in the pavement deterioration functions, Equations 3 and 4. Note that when S = 0, C(S) = C_0. With this function it is possible to model user costs as rising at a slow, roughly linear pace between about 5 PSI and 2 PSI and then rising rapidly thereafter.

Rehabilitation Cost

In this section a relationship for I as a function of S^{*-} is developed. The approach used combines principles from the AASHTO Interim Guide for Design of Pavement Structures (3) along with a hypothesized relationship between the pavement's surface layer coefficient and the present serviceability index. It should be understood from the outset that the specific function for $I(S^{*-})$ developed herein is primarily a proposal or suggestion and should not be construed as representing the definitive relationship between overlay cost and current pavement condition. It is intended more as an example used to illustrate the broad capabilities of this model.

Although overlays have been chosen to illustrate the mathematical formulation of the model and its application to example solutions, other classes of rehabilitation (e.g., recycling, major sublayer stabilization, surface restoration) could also be treated. Different types of rehabilitation would be represented in the model through the particular form of the pavement deterioration function assigned to Stage 2 in Figure 2, the values assigned to the parameters of this function, and the unit costs input.

Two rehabilitation cases are developed as examples in this paper: flexible overlays of flexible pavements and flexible overlays of rigid pavements. Because the functions relating design procedures (and hence investment cost) to pavement condition are somewhat different for these cases, the appropriate equations will be developed for each case separately. The general methodology is to determine an overlay thickness based on structural condition and future traffic predictions and then to determine cost from the required overlay thickness.

FLEXIBLE OVERLAYS OF FLEXIBLE PAVEMENTS

The first task is to find the overlay thickness that is needed to restore the pavement's structural number to its original design value in the as-constructed state. Assuming that the flexible pavement consists of a subbase, a base, and a surface layer, the overlaid pavement will have four layers: subbase, base, original surface, and overlaid surface. Equating the structural number of the original pavement to that of the newly overlaid pavement:

$$a_1 D_1 + a_2 D_2 + a_3 D_3 = a_1 D_1 + a_2 D_2 + a_3' D_3 + a_4 D_4 \quad (9)$$

where

a_1, a_2, a_3, a_4 = layer coefficient for subbase, base, original surface, and overlaid surface layers, respectively;

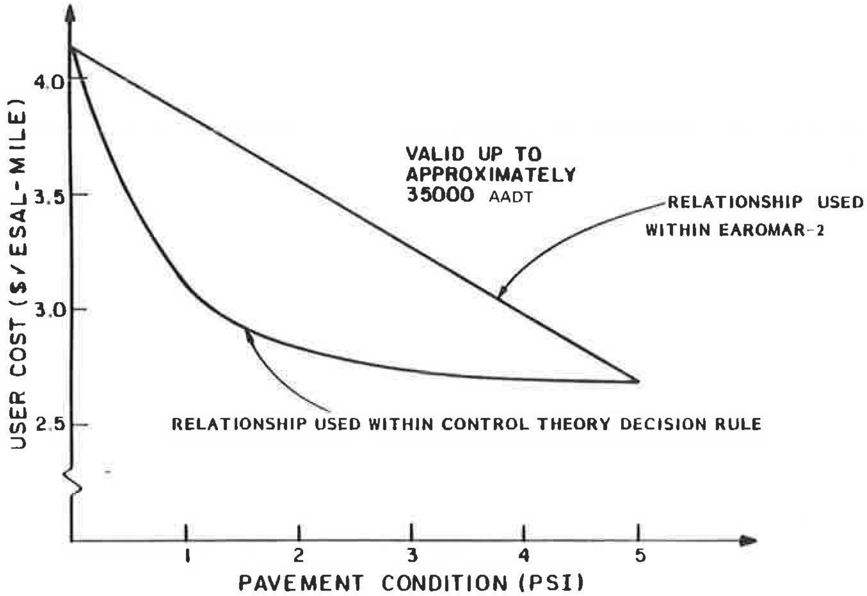

FIGURE 5 Comparison of nonlinear C(S) function to linear C(S) function.

D_1, D_2, D_3, D_4 = thickness of subbase, base, original surface, and overlaid surface layers, respectively; and

a_3' = decayed value of layer coefficient corresponding to current condition of original surface layer.

Equation 9 implies that for design purposes the strength of the existing pavement due to be overlaid is reduced from its as-constructed value. In accordance with AASHTO (3,7) this is represented by lowering the value of the original surface layer coefficient from a_3 to a_3'. Assuming that the layer coefficients and thicknesses of the subbase and base do not change over time, Equation 9 yields

$$D_4 = [(a_3 - a_3')/a_4]\ D_3 \tag{10}$$

Recognizing that the layer coefficient of the overlay usually equals the layer coefficient of the original surface layers, Equation 10 becomes

$$D_4 = D_3\ [1 - (a_3'/a_3)] \tag{11}$$

The fact that the surface layer coefficient of a pavement due to be overlaid is reduced from its as-constructed value can be used to establish a relationship between the surface layer coefficient and the PSI. It is hypothesized that the surface layer coefficient experiences an exponential decay during pavement life. This assumption is similar to the research reported for the EAROMAR-2 program (4). Mathematically,

$$a_3' = a_3 e^{\lambda[S_1(t) - S_0]} \tag{12}$$

Combining Equations 11 and 12 with Equation 3 results in

$$D_4^S = D_3 \left[1 - e^{\lambda k_1 (1 - e^{\beta_1 t^*})}\right] \tag{13}$$

where D_4^S represents the overlay thickness needed to restore the pavement's structural number to its original design value.

The overlay thickness specified by Equation 13, however, does not account for the additional traffic loading beyond the original design level that the pavement might experience during the next t years. This additional thickness, to be labeled D_4^Q, can be accounted for by taking the difference between the structural number as specified by AASHTO for the design of a flexible pavement (with t* as the beginning of the design period) and the structural number specified is Equation 9 and then dividing this difference by the overlay surface coefficient. The resulting differential thickness due to traffic growth is

$$D_4^Q = (1/a_4)\ [SN_{TOT} - a_1 D_1 - a_2 D_2 - a_4 D_3] \tag{14}$$

SN_{TOT} can be determined using the AASHTO design procedure for flexible pavement structures (3). This procedure provides a relationship between weighted structural number and total equivalent 18-kip single axle load applications (the model's measure of traffic volume). The relationship between axle loads over a 20-year life, Q_{20}, and structural number, SN, can be approximated as follows (2):

$$SN = (Q_{20}/310)^{1/6} \tag{15}$$

with $R^2 = 0.99$.

The value of Q_{20} can be obtained by integrating Equation 5 from t* to some time τ:

$$Q_\tau = \int_{t^*}^{t^*+\tau} [A_2 t - B_2 e^{\beta_2 (t - t^*)} + C_2]\ dt \tag{16}$$

where (for brevity) $A_i \equiv q_0\ (ak_i + aS_A + b)$, $B_i \equiv q_0\ ak_i/\beta_i$, and $C_2 \equiv A_1 t^* - B_1 \exp\ (\beta_1 t^*) + C_1 + B_2 - A_2 t^*$. The resulting solution is

$$Q_\tau = \tau\left\{A_2 [t^* + (1/2)\tau] + C_2\right\} - (B_2/\beta_2)(e^{\tau\beta_2} - 1) \tag{17}$$

where τ can be set equal to 20 years (or any other design life). The required thickness for a flexible overlay of a flexible pavement may be written as

$$D_4 = D_3 \left[1 - e^{\lambda k_1 \left(1 - e^{\beta_1 t^*}\right)}\right] + (1/a_4) \left\{[(Q_\tau/310)^{1/6}] - a_1 D_1 - a_2 D_2 - a_4 D_3\right\} \tag{18}$$

where Q_T is given by Equation 17; the first term of Equation 18 represents D_4^S and the second term represents D_4^Q. Now the cost of the overlay can be found by multiplying the thickness by a factor of cost per unit of thickness:

$$I = \sigma_F(D_4) \tag{19}$$

As a final note, it is possible for Equation 14 to give a negative value for D_4^Q if the original pavement design had a higher structural number than that required by the AASHTO design method applied at t*. (Such a case might arise, for example, if the original pavement design procedure differed from that of AASHTO.) In this case D_4^Q should simply be set equal to zero. Moreover, the overlay thickness determined by Equation 18 will be subject to a minimum thickness constraint required by construction procedures.

FLEXIBLE OVERLAYS OF RIGID PAVEMENTS

A rigid pavement given a flexible overlay becomes a composite pavement. This analysis treats a composite pavement as a flexible pavement with a relatively strong base (the former rigid surface layer). In this case, however, a single overlay thickness, which accounts for both restoring the pavement to its original strength and allowing for future traffic loadings, is determined.

For the composite pavement,

$$SN_{TOT} = a_s D_s + a'_R D_R + a_F(D_0) \tag{20}$$

where

a_s, a'_r, a_F = layer coefficient for subbase, rigid slab, and flexible overlay, respectively;
D_s, D_R = thickness of subbase and rigid slab, respectively; and
D_0 = required flexible overlay thickness.

In accordance with AASHTO (7) the layer coefficient of the rigid slab about to be overlaid is reduced from its original value. As in the previous section, this can be used to hypothesize a relationship between the layer coefficient and surface condition:

$$a'_R = a_R\, e^{\lambda[S_1(t) - S_0]} \tag{21}$$

where in general λ for flexible overlays of rigid pavements will not equal λ for flexible overlays of flexible pavements. In addition, Equation 15 can be used to determine the required structural number from the AASHTO procedure. Therefore, substituting Equations 15 and 21 along with 3 into Equation 20 and solving for D_0 yields

$$D_0 = (1/a_F)\left\{[(Q_T/310)^{1/6}] - a_s D_s - a_R D_R e^{\lambda k_1(1-e^{\beta_1 t^*})}\right\} \tag{22}$$

where Q_T is given by Equation 18.

The cost of the overlay can be found by multiplying the thickness by a cost factor (σ_R) as defined in the previous section. Once again, Equation 22 may produce a negative value if the original rigid pavement was overdesigned compared to AASHTO criteria. A minimum overlay thickness may be used to override this event.

Salvage Value

The salvage value in this model is assumed to depend on an infinite stream of user benefits beginning at T. [This assumption was originally made by Fernandez (1).] Mathematically this can be expressed as

$$\psi(T) = \int_T^{\infty} -G_2(t)q(T)\exp(-\rho t)dt \tag{23}$$

where $-G_2(t)$ is defined as $\{U-[C(S,q) + (\partial C/\partial q)\, q_2(t)]\}$ and represents net benefits per user during Stage 2. The term q(T) represents the number of users at terminal time T; for simplicity it is assumed that it becomes constant at this time. Multiplying the infinite stream of per user net benefits by the traffic volume at T yields the infinite stream of total user net benefits.

Resulting Decision Rule

The complete decision rules for flexible and for rigid pavements are as follows:

For Flexible Overlays of Flexible Pavements

$$\begin{aligned}\rho\sigma(D_4) = {} & \mu C_0 \left[e^{-\omega(k_1 + S_0 - k_1 e^{\beta_1 t^*})} - e^{-\omega S_A}\right] x\,(A_1 t^* - B_1 e^{\beta_1 t^*} + C_1) \\ & + \left[S_A - S_0 + k_1 (e^{\beta_1 t^*} - 1)\right] \pi(t^*) a q_0 \\ & + \sigma D_3 \lambda \beta_1 k_1 e^{\lambda k_1 (1-e^{\beta_1 t^*})} + \beta_1 t^* \end{aligned} \tag{24}$$

where $\pi(t^*)$ is given by Equation 26 and D_4 is given by Equation 18.

For Flexible Overlays of Rigid Pavements

$$\begin{aligned}\rho\sigma(D_0) = {} & \mu C_0 \left[e^{-\omega\left(k_1 + S_0 - k_1 e^{\beta_1 t^*}\right)} - e^{-\omega S_A}\right] x\,(A_1 t^* - B_1 e^{\beta_1 t^*} + C_1) \\ & + \left[S_A - S_0 + k_1 (e^{\beta_1 t^*} - 1)\right] \pi(t^*) a q_0 \\ & + (a_R/a_F)\, \sigma D_R \lambda \beta_1 k_1 e^{\lambda k_1 (1-e^{\beta_1 t^*}) + \beta_1 t^*} \end{aligned} \tag{25}$$

where $\pi(t^*)$ is given by Equation 26 and D_0 is given by Equation 22.

$$\begin{aligned}\pi(t^*) = (U/\rho) - \Big\{ & -[\mu C_0/(\beta_2 \omega k_2 - \rho)]e^{-\omega S_A} + [(1-\mu)C_0/\rho] \\ & + (n+1)\delta[1/\rho\,(A_2 t^* + C_2)^n] + (nA_2/\rho^2)(A_2 t^* + C_2)^{n-1} \\ & + [n(n-1)A_2^2/\rho^3](A_2 t^* + C_2)^{n-2} \\ & + [n(n-1)(n-2)A_2^3/\rho^4](A_2 t^* + C_2)^{n-3} \\ & + [n(n-1)(n-2)(n-3)A_2^4/\rho^5](A_2 t^* + C_2)^{n-4} \\ & + [n(n-1)(n-2)(n-3)(n-4)A_2^5/\rho^6](A_2 t^* + C_2)^{n-5} + (n!A_2^6/\rho^7) \Big\} \end{aligned} \tag{26}$$

Although Equations 24 and 25 appear complex, their interpretation yields some intuitive insights into the structure of the solution.

The rules are marginal rules, balancing marginal costs (on the left side of the equations) and marginal benefits (on the right side of the equations). The term on the left denote the capitalized costs of undertaking the rehabilitation investment: in this case, the overlay of either flexible or rigid pavement. The first term on the right side of Equations 24 and 25, respectively, denotes the benefits of the rehabilitation accruing to the traffic stream, resulting from reductions in user costs due to the improved quality of the pavement surface. The second term in each equation quantifies the benefits of attracting additional traffic (and thereby providing the advantages of transportation to more users) because of the improved quality of the pavement. (Of course, additional traffic also causes increased rates of pavement deterioration and of congestion; these effects can be captured in the deterioration and user cost equations discussed earlier. Also note that if the variable "a" defined in Equation 5 is zero, this "generated traffic" effect is eliminated.) The third term in each equation captures the monetary benefit associated with preservation of investment (i.e., if rehabilitation is performed earlier, more substantial rehabilitation is avoided later). This term, in effect, justifies the avoidance of "deferred maintenance."

APPLICATIONS TO EXAMPLES

General Information

The decision rules in Equations 24 and 25 were applied to a series of examples of flexible and rigid pavement rehabilitation. The general approach was to define, for each pavement type, five arbitrary designs of different strengths: F1 through F5 for flexible pavements and R1 through R5 for rigid pavements. Each pavement design was subjected to four different traffic levels, corresponding to 5,000 AADT, 15,000 AADT, 25,000 AADT, and 35,000 AADT at the start of the analysis period (i.e., before growth). For each combination of pavement design and traffic level, the optimal rehabilitation time, t^*, was computed by solving either Equation 24 or Equation 25. The results were plotted to assess the general trends of the solution and to provide an easy way for engineers and administrators to apply these results in practice.

Given the design of this approach, not all the pavement-traffic combinations represent desirable situations. For example, some combinations impose heavy traffic on weak pavements, and others test light traffic on strong pavements. Nevertheless, including such combinations along with the more closely matched traffic-pavement design pairs has two advantages. First, it allows development of the solution function over a wide domain of traffic and pavement design possibilities and investigation of the behavior of the solution at the boundaries of typical situations. Second, it recognizes that, in their focus on existing pavements, rehabilitation decisions are different from those of design and new construction. (Refer to the several points at the beginning of this paper.) It is plausible that a pavement, once built, will be subjected to traffic levels much lighter or much heavier than that for which it was designed. Therefore, it should be possible to consider at least the possibility of some unforeseen combinations of design thickness and traffic.

Description of Examples

The numerical examples involved a two-lane, one-directional roadway with a design speed limit of 70 mph. The environmental region simulated was the northeastern United States.

The structural designs of the five rigid pavements and five flexible pavements tested are given in Tables 1 and 2. The portland cement concrete pavement is a plain jointed slab over a granular subbase; the asphalt pavement consists of an asphalt concrete surfacing over a granular base and subbase.

TABLE 1 Characteristics of Rigid Pavement Designs

Design	Subbase Thickness (in.)	Surface Thickness (in.)
R1	6	8
R2	6	9
R3	8	10
R4	9	11
R5	10	12
Layer coefficient	0.11	0.50

TABLE 2 Characteristics of Flexible Pavement Designs

Design	Subbase Thickness (in.)	Base Thickness (in.)	Surface Thickness (in.)
F1	12	6	4
F2	16	9	6
F3	24	12	8
F4	24	12	10
F5	24	12	12
Layer coefficient	0.11	0.14	0.44

In developing the examples, values had to be assigned to the parameters in Equations 24 and 25. These parameters can be grouped into subsets corresponding to the model's basic functional categories: pavement deterioration, $S(t)$; traffic demand, $q(t)$; user cost, $C(S,q)$; and investment cost, $I(S^{*-})$. To obtain these calibrations, the EAROMAR-2 simulation program was used to quantify general behavioral trends for these categories, using the same pavement designs, traffic loads, environmental conditions, maintenance policies, and other factors defined for the examples. From these results, specific values were assigned to the parameters in the model's detailed equations. Again, overlays were used to illustrate model application; however, other types of rehabilitation can also be represented under this approach.

The list of calibrated values is extensive and is presented elsewhere (2). Following are summary data that highlight key elements of the examples.

Pavement Deterioration

Values of k and β were determined to represent pavement deterioration for the various combinations of design and traffic level. To illustrate the results obtained, Figure 6 shows the predicted deterioration of the five rigid pavements tested for an initial traffic of 25,000 ADT; Figure 7 shows corresponding curves for flexible pavement.

The curves in Figures 6 and 7 correspond to Stage 1 (before the overlay), and rates of deterioration are quantified by β_1 and k_1 in Equation 3. After the overlay, the rigid pavement was treated as a compos-

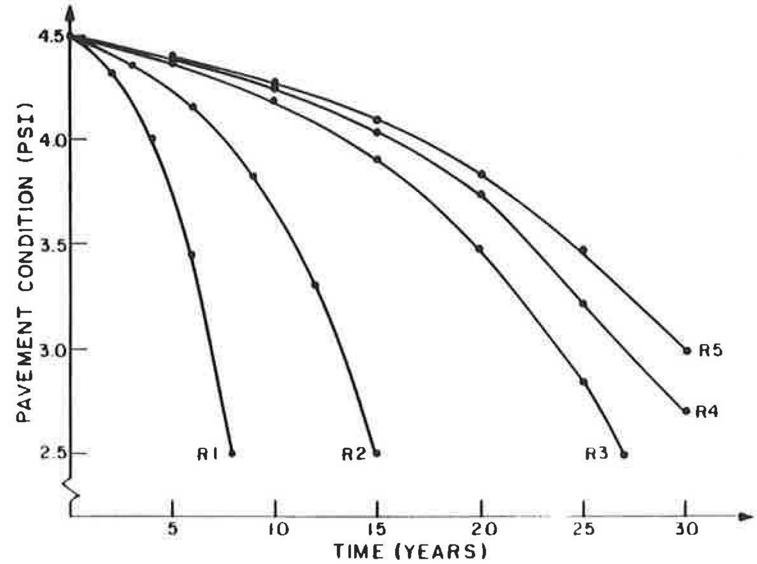

FIGURE 6 Rigid pavement deterioration curves for 25,000 AADT.

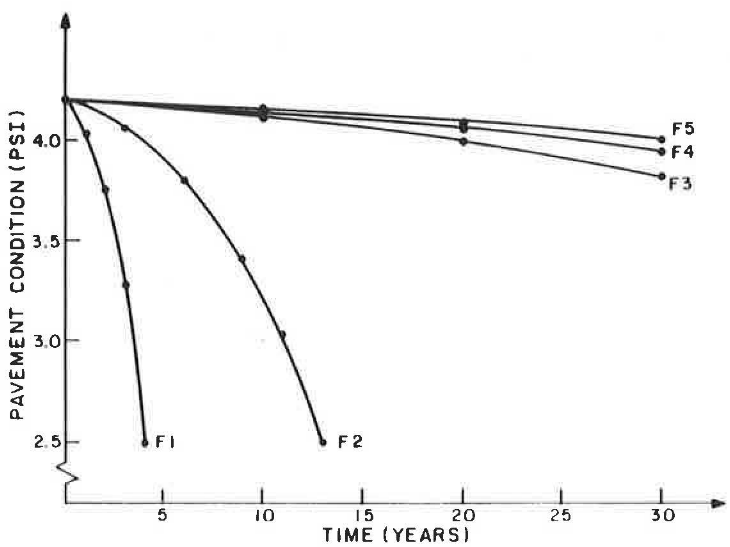

FIGURE 7 Flexible pavement deterioration curves for 25,000 AADT.

ite pavement, as described earlier; therefore, new values were determined for k_2 and β_2 in Equation 4, for each composite design C1 through C5 (corresponding to R1 through R5 after overlay), and for each traffic level (2). Also note that for 25,000 ADT, flexible designs F3 through F5 are simulated to act as premium pavements, showing only slight loss in PSI over a 30-year period. (Again, recall that a range of arbitrary designs was selected for the examples.)

Traffic Demand

Four cases of traffic demand were defined as discussed earlier, ranging from 5,000 to 35,000 AADT at the start of the analysis (t = 0). Because the model represents traffic in 18-kip equivalent single axle loads (ESALs), the AADT values were converted to ESALs. For the particular traffic stream simulated in EAROMAR-2, the conversion factor used was 0.1099 ESAL per vehicle. (Notwithstanding this conversion, traffic will continue to be described by AADT in this paper to retain clarity.)

Traffic growth is represented by the parameters a and b in Equation 5. Either positive or negative values for a and b are permissible; this example assumed a = 0.001 [or 0.1 percent per year per unit change in pavement condition S(t)] and b = 0.01 (or 1 percent per year).

User Costs

User costs were separated into a congestion effect and a pavement-related effect. The congestion functions simulated by EAROMAR-2 and used in the decision rule are shown in Figure 8.

The important point to observe from Figure 8 is that traffic volume itself has virtually no effect on user cost up to a certain threshold where congestion begins. As congestion increases, user costs

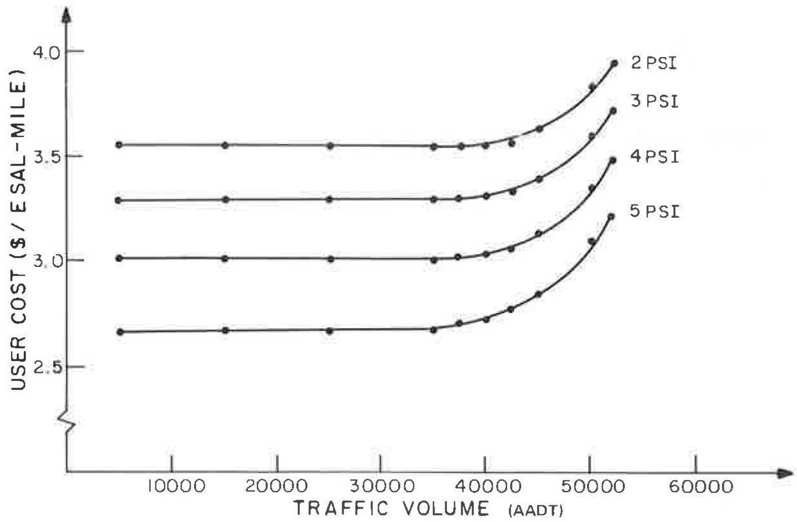

FIGURE 8 User cost versus traffic volume.

rise very quickly, due primarily to the dramatic rise in travel time costs. For the case of a two-lane, one-directional high-speed roadway, the EAROMAR-2 simulations indicate that user costs begin to be affected by traffic volume at approximately 37,500 AADT. The value of 37,500 AADT is henceforth referred to as the congestion threshold. The curves in Figure 8 are captured by the following relationship:

$$C(q) = \begin{cases} \text{constant}, 0 < q < q' \\ \delta(q-q)^n + K, q' \leq q < q_{max} \end{cases} \tag{27}$$

where q' is the congestion threshold.

The relationship of user costs to pavement condition was the subject of some review, as discussed earlier with respect to Figure 5. Based on findings of Zaniewski (5) and Ross (6), values were estimated for the parameters given in Equation 8, resulting in the following relationship used in the model:

$$C(S) = 1.4528 \exp(-1.25\ S) + 2.6972 \tag{28}$$

User Utility

User utility may be regarded as equal to the largest cost that a user will tolerate and still choose to use the roadway. In other words, it must be worth at least this much for a user to occupy the roadway; otherwise, he or she simply would not use it. This is true even for captive traffic that suffers high cost while occupying a roadway only because there exist no alternative links. This traffic still chooses to use the road and hence the utility to these users of doing so must be at least equal to the cost that they absorb.

In the development of the rehabilitation investment decision model, it has been assumed that utility is constant and the same for each user (a simplifying assumption to reduce mathematical complexity and field calibration requirements). For the current example, it was assumed that user utility is equal to the user cost associated with 1.0 PSI and 50,000 AADT. For this traffic level, a nonlinear C(S) curve similar to the one in Figure 5 indicates that this value is about $3.45/ESAL·mile, or roughly $0.28/vehicle·mile.

Overlay Costs

Building Construction Cost Data 1981 (8) indicates that $\$1.62/yd^2 \cdot in.$ is a representative cost estimate for placing a bituminous wearing course for asphalt priced at $19.95 per ton. (The figure of $\$1.62/yd^2 \cdot in.$ can be adjusted to reflect price changes.) This estimate includes materials, installation, and contractor's overhead and profit.

The 1-mi length of roadway assumed for the case study encompasses 23,466.67 yd^2/mile; the cost of placing a new bituminous wearing course therefore becomes approximately $38,000/in.·mile. However, this does not represent the only cost associated with the investment. There are also base preparation, mobilization, and line painting, as well as the public highway department's design, inspection, and general overhead costs. To account for these costs, 16 percent has been added to the estimate. The investment cost parameter, labeled σ, used in the case studies therefore equals $44,100/in.·mile; thus, a 2-in. overlay would cost $88,200/mile, and a 3-in. overlay would cost $132,300/mile.

To represent the variability in overlay costs over time, λ in Equations 21 and 22 was estimated to equal 0.2554. A minimum overlay thickness of 2 in. was specified.

Discount Rate

A discount rate of 7 percent was used (assuming constant dollar estimates excluding inflation).

Optimal Rehabilitation Times

The solutions for the optimal time of rehabilitation investment are most easily presented in graphic form. The solutions for rigid pavements developed for the case study are shown in Figures 9 and 10. These two figures represent the same set of solutions plotted in different ways. Similarly, Figures 11 and 12 show the solutions for flexible pavements. The trends portrayed by these curves may be used in making a number of decisions.

The most direct application is in the programming of rehabilitation expenditures. For a given pavement and traffic, the optimal time to rehabilitate may be

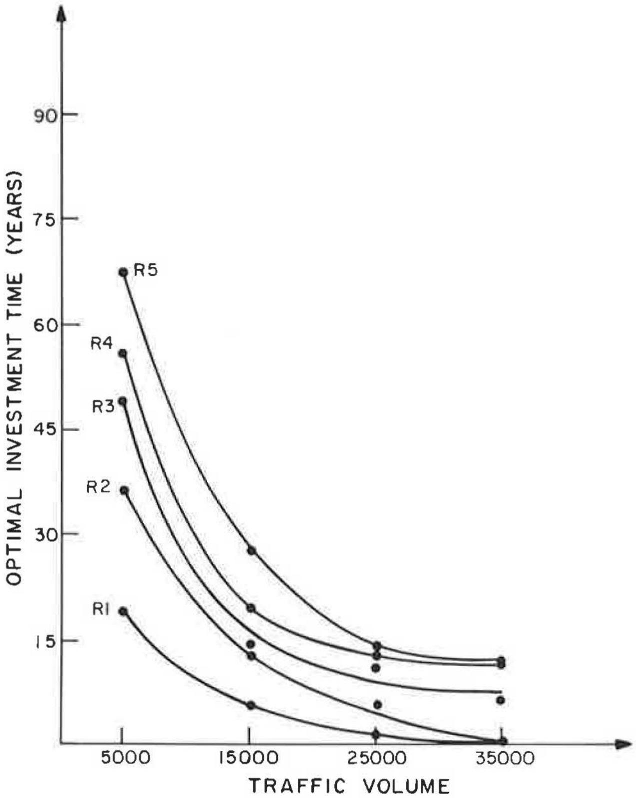

FIGURE 9. Optimal investment time versus traffic volume for variable traffic and variable investment cost (rigid pavement); note: traffic volumes represent initial values.

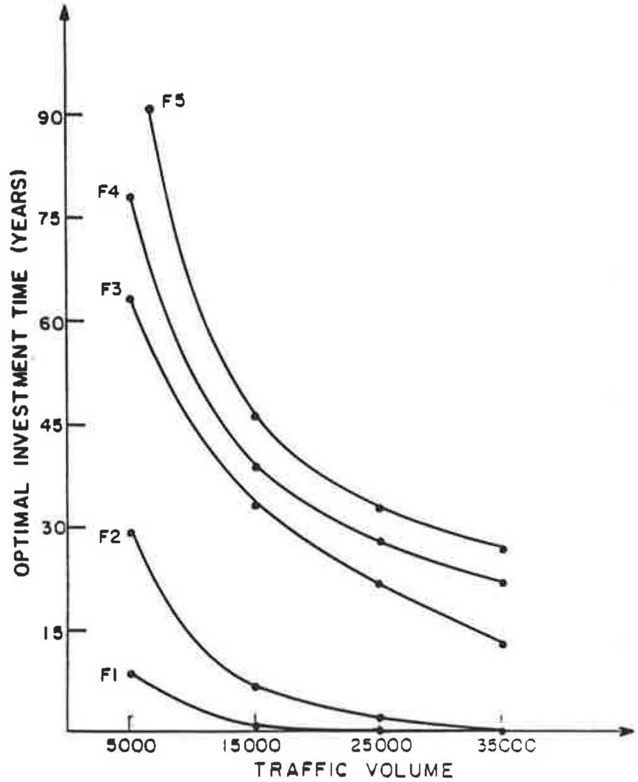

FIGURE 11 Optimal investment time versus traffic volume for variable traffic and variable investment cost (flexible pavement); note: traffic volumes represent initial values.

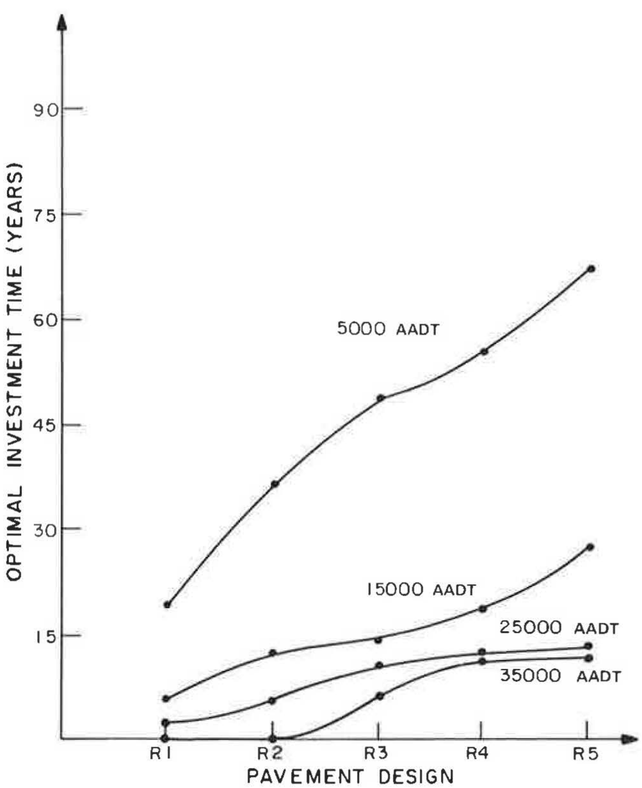

FIGURE 10 Optimal investment time versus pavement design for variable traffic and variable investment cost (rigid pavement); note: traffic volumes represent initial values.

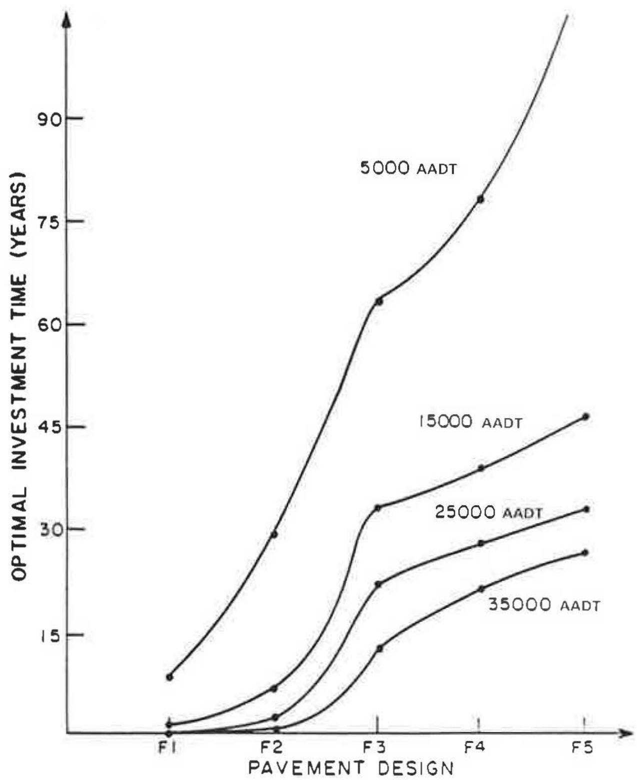

FIGURE 12 Optimal investment time versus pavement design for variable traffic and variable investment cost (flexible pavement); note: traffic volumes represent initial values.

determined and used as the basis for scheduling work and allocating funds. Either Figure 9 or Figure 10 for rigid pavements, or Figure 11 or Figure 12 for flexible pavements may be used for this purpose. It is also possible to organize a series of such calculations for a pavement network and to develop rehabilitation programs on either an open-ended or a budget-constrained basis.

The second application is in the evaluation of design-rehabilitation trade-offs. For example, design procedures such as AASHTO's (3) are fixed by an assumed 20-year pavement life. However, Figures 9 and 11 indicate that, for a given traffic projection, a number of designs and design lives are possible. For example, assuming an initial traffic of 15,000 AADT for rigid pavements, designs with optimal rehabilitation times of about 15 years (e.g., R2 or R3) or stronger pavements with longer optimal lifetimes (e.g., 19 years for R4, or 28 years for R5) could be selected. Note, however, that Figures 9 and 11 simply indicate the optimal investment time for a given set of circumstances (in this case, pavement design and traffic level); they do not indicate which design-rehabilitation combination has the lowest life-cycle cost. This cost information can be obtained, however, by evaluating the objective function, Equation 6, at t* for each pavement design. (In this case, the costs of pavement initial construction must be included in the objective function.)

The third application is the conduct of sensitivity analyses of pavement design and rehabilitation with respect to traffic volume. For example, consider the curves for the five rigid designs in Figure 10, and assume that the best available traffic load projection for a pavement corresponds to 25,000 AADT, but is subject to considerable uncertainty. Observe from Figure 10 that, at this level of traffic, designs R1 and R2 are relatively sensitive to changes in traffic, whereas R3 through R5 are less so. Again, only sensitivity is indicated by the curves in Figures 10 and 12; to understand the cost impacts, the objective function, Equation 6, at the solution t* must be evaluated.

One remaining point that is important to reiterate is that the solutions defined by Equations 24 and 25, and illustrated by Figures 9-12, are based on both economic and technical criteria, as opposed to the purely technical criteria traditionally applied to pavements. For example, a common standard derived from the AASHTO Road Test is that high-type pavements be overlaid at a PSI of 2.5. However, in the results indicated by Figures 9-12, the pavements, for the most part, were overlaid at PSI values higher than 2.5. This trend appears to be consistent with results of a survey of highway departments conducted by AASHTO (Summary of Selected State Practices Collected in 1980 Through AASHTO for the Truck Size and Weight Study, Section 161, Surface Transportation Assistance Act of 1978, memorandum, Federal Highway Administration, December 1982), showing considerable variability in the actual threshold for overlays used by different states. (Bear in mind also that the model is predicting a desired, not the actual, threshold.) Moreover, note that of the three cases defined by Balta (2), the case reported herein is the one most likely to drive pavement rehabilitation earlier, because its assumptions favor to a greater degree the benefits to both highway agency and road users of a high-quality pavement that is sustained by more frequent overlays. The adoption of one of the alternative assumptions investigated by Balta (i.e., the case of constant traffic over time, or the case of constant rehabilitation cost over time) would tend to defer the optimal rehabilitation time.

CONCLUSION

This paper began with the premise that maintenance and rehabilitation are inherently different from new construction because they involve an existing facility and require an understanding of facility performance as opposed to design. Moreover, the optimization of maintenance and rehabilitation policy is difficult because of the different options available in the choice and the timing of activities because pavement performance and related costs change over time, with no definitive point of failure. Methods of evaluating maintenance and rehabilitation policy cannot be based solely on technical grounds but must also include life-cycle costs and benefits.

As an example of one promising avenue, pavement rehabilitation, has been investigated and an analytic procedure to determine the optimal time to overlay for both flexible and rigid pavements has been developed and investigated. The procedure is based on principles of dynamic control theory and yields results that can be organized within sets of curves that are easy to understand and use. The control theory approach is predicated on an objective function rooted in engineering and economic principles and subject to constraints representing the detailed technical performance of the pavement-traffic interaction.

A series of examples has been presented to illustrate the nature of the solution and demonstrate the practicality of the results. Data for these examples were obtained from the EAROMAR-2 simulation model. However, to be a truly effective tool, the control theory solution should be calibrated by relationships validated in the field. This is true for all categories of parameters identified earlier in the paper but is especially true for user costs, which play a strong role in driving the solution of t*, the optimal time to rehabilitation.

Several potential applications of the control theory solutions have been discussed, including the programming of pavement rehabilitation, the investigation of design-rehabilitation trade-offs, and sensitivity analyses. More generally, the control theory solution for rehabilitation, coupled with an analogous solution for routine maintenance, could play an important role in many aspects of pavement management.

ACKNOWLEDGMENTS

The work reported in this paper was sponsored by the U.S. Department of Transportation through the Federal Highway Administration. The authors would like to especially thank William Kenis, Program Monitor, for his cooperation and encouragement throughout this study and for his assistance in providing data for the cases. The authors would also like to recognize and gratefully acknowledge the contributions of Mark Thurston for his research effort at M.I.T., particularly in the formulation of the case study data, execution of the simulation model, and analyses of results.

REFERENCES

1. J.E. Fernandez Larranaga. Optimal Dynamic Investment Policies for Facilities: The Transportation Case. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, 1978.
2. W.S. Balta. Optimal Investment Policies for Maintenance and Rehabilitation of Highway Pavements.

M.S. thesis. Massachusetts Institute of Technology, Cambridge, 1984.
3. AASHTO Interim Guide for Design of Pavement Structures. American Association of State Highway and Transportation Officials, Washington, D.C., 1972 (revised ed. 1981).
4. M.J. Markow and B.D. Brademeyer. EAROMAR Version 2: Final Technical Report. Report FHWA/RD-82/086. FHWA, U.S. Department of Transportation, April 1984.
5. J.P. Zaniewski. Fuel Consumption Related to Roadway Characteristics. Texas Research and Development Foundation, Jan. 1983.
6. F.R. Ross. The Effect of Pavement Roughness on Vehicle Fuel Consumption. Materials Section, Division of Highways and Transportation Facilities, Wisconsin Department of Transportation, Madison, June 1981.
7. The AASHO Road Test: Proceedings of a Conference Held May 16-18, 1962, St. Louis, Mo. HRB Special Report 73. HRB, National Research Council, Washington, D.C., 1962.
8. Building Construction Cost Data 1981, 39th annual ed. Robert Snow Means Company, Inc., 1982.

Publication of this paper sponsored by Committee on Maintenance and Operations Management.

Pavement Routine Maintenance Cost Prediction Models

ESSAM A. SHARAF, KUMARES C. SINHA, and VIRGIL L. ANDERSON

ABSTRACT

In this paper a methodology is presented for using the available data on pavement routine maintenance from the Indiana Department of Highways (IDOH) to develop models relating the cost of pavement routine maintenance to pavement system characteristics on a network level. The results showed that total pavement routine maintenance costs are affected by traffic level and by climatic zone (weather effect). Furthermore, the analysis of costs of individual activities showed that the extent of patching work (amount of pothole repair that is done after winter) is negatively correlated to the amount of sealing activity that takes place before winter. The implication of this result is that a higher level of service (fewer potholes) may be achieved by increasing sealing activity.

One of the common shortcomings of most current highway maintenance management systems (MMS) is that they are primarily designed for managing available resources (labor, materials, and equipment) and not geared to managing pavement facilities (1). The focus in this paper is on the use of available maintenance management data to provide information that can be directly employed in highway pavement management. In particular, models for pavement routine maintenance costs, which can be effectively used in preparing annual maintenance programs as well as in making decisions about resurfacing and rehabilitation, particularly on a network level, were developed.

As is the case in many other states, the maintenance management system in Indiana is designed for resource management and the necessary data are recorded on an aggregated unit representing a subdistrict. However, other pavement-related information is recorded on the basis of a contract. On the average, a subdistrict may include more than 100 contracts. The nonconformity between the maintenance data and the pavement data makes it difficult to use MMS information effectively in pavement management. For the purpose of this study a system was developed to represent all available data in terms of a highway section that was defined as the part of a highway within a county limit. This system allowed the maximum use of both the MMS data and the pavement management data as a unified information base. The