Dynamic Model for Scheduling Maintenance of Transportation Facilities

KIYOAKI MURAKAMI and MARK A. TURNQUIST

ABSTRACT

The optimum scheduling of maintenance for transportation facilities is addressed. The problem is described as a multiple-period resource allocation problem with constraints on both resource availability and state and decision variables. The problem is formulated as a nonlinear optimization problem and solved by using the generalized reduced gradient method. The model uses recursive formulas similar to those used in dynamic programming in order to calculate the partial derivatives of the objective function. The model is applied to an example based, in part, on actual data provided by the Japanese National Railways. Several tests are made to show the performance of the model, and the results are compared with those of two alternative solutions. The results show the usefulness of the model in a wide variety of applications and its superiority to the alternative solutions examined.

Maintenance of facilities is important in all transportation modes. Maintenance consumes a sizable, and increasing, fraction of total operating expenditures. For example, the 1982 Highway Cost Allocation Study (1) estimated 39 percent of all highway expenditures in 1985 would be for resurfacing, restoring, rehabilitating, and reconstructing (4R) work compared to 25 percent in 1978. In 1982 U.S. railroads spent $5.2 billion on maintenance of way and structures (19.7 percent of total operating costs), up from $3.5 billion (17.7 percent of total operating costs) 5 years earlier (2).

An important characteristic of many transportation facilities is that their condition declines nonlinearly over time. Figure 1 shows this point, using as an example the pavement present serviceability index (PSI) described by AASHTO (3). Because of this characteristic, the timing of maintenance expenditures is important. The marginal effectiveness of an additional dollar spent for maintaining a given facility depends greatly on the condition of that facility when the maintenance is performed. This characteristic is important not only for highway systems but for railroad and waterway facilities as well.

In this paper a general method for planning maintenance expenditures over multiple time periods is developed. The model includes the deterioration characteristics of the facilities under study, the dependence of maintenance effectiveness on current

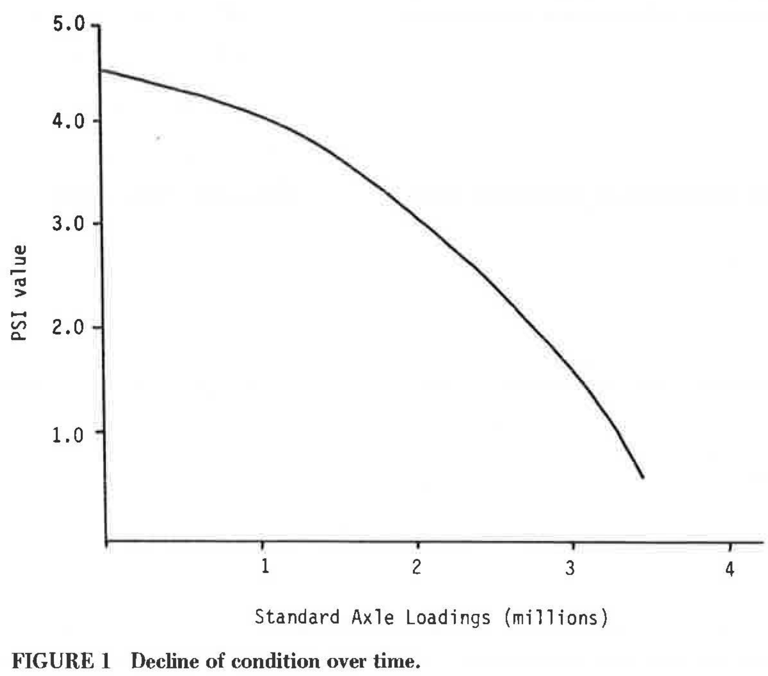

FIGURE 1 Decline of condition over time.

system state, and the potential variation in available maintenance resources across time periods.

The model is formulated as a nonlinear optimization problem that is readily solvable by standard nonlinear programming (NLP) computer codes. The model determines the allocation of available maintenance resources to each of several facilities (or sections of a single facility) in each time period, which maximizes an overall measure of system condition.

In the following section of the paper the structure of the model is described in greater detail. The third section presents an example application in railroad track maintenance and develops the mathematical formulation. The fourth section describes a solution method. Illustrative calculations for an example problem emphasize the relative advantages of the method. The last section presents conclusions and emphasizes the potential applicability of the model to highway and waterway problems, as well as to railroads.

MODEL DESCRIPTION

A basic premise of this model is that maintenance is a regular periodic activity and that the decisions that must be made are the amount of resources to be allocated to maintenance of a particular facility (or section of a facility) in each of a number of periods. Suppose that at the beginning of period j the state (or condition) of a specific section is given. In the absence of maintenance, that state will deteriorate during period j as a result of use and weather. The rate of deterioration is a basic element in this model.

This deterioration can be overcome by assigning maintenance resources to the section. The effectiveness of a particular set of resources in upgrading facility condition will depend on the current state of the facility, its structural characteristics (which determine "maintenance effectiveness"), and the available working time during the period (determined by traffic levels and climate).

All of these basic elements of the problem are combined in what is denoted the transition function, which specifies the facility state or condition at the beginning of period j+1 as a function of the state at the beginning of period j, the deterioration rate during period j, the maintenance resources allocated during period j, the available working hours during period j, and parameters determining the effectiveness of those resources. These elements are shown in Figure 2.

The assignment of maintenance resources is subject to the following constraints:

1. The available resources are limited;
2. The facility state for each section must stay above certain minimum standards; and
3. The resources to be assigned to a certain section may be bounded both from below and from above because of availability of facilities or operators or because of management policy.

One of the most important aspects of resource assignment is that it must be treated as a dynamic problem. In general, the three major parameters (deterioration rate, maintenance effectiveness, and available working hours) are not constant with location or time. For example, in track maintenance planning for railroads, available tamping hours vary from location to location and time to time because the actual work is performed during train intervals that are longer than a certain number of minutes. Therefore, in sections where the train frequency is high, the available tamping hours are extremely limited so the effectiveness of each tamping machine is reduced. Moreover, it is physically impossible to perform the work in heavy rain or when there is snow on the track. Because of the danger of rail buckle, the work is also restricted when the rail temperature is high.

In addition to the parameters, the constraints may be dependent on location or time. The bound on the facility state and the upper and lower bounds on the decision variables may differ from location to location, although they are usually considered constant with time. The total available resources also will frequently vary from one time period to another.

In this study the facilities are divided into several sections and the time into several periods such that the facility state within a section can be considered homogeneous. Thus the problem is how many resources should be assigned to each section in each period in order to achieve some overall objectives, subject to a set of constraints.

EXAMPLE APPLICATION

As an example of the general ideas expressed in the previous section, a specific application to railroad track maintenance planning is considered.

The track irregularity index (denoted the P-index) is used by Japanese National Railways (JNR) for the purpose of track state control. It shows the percentage of track irregularities within a certain

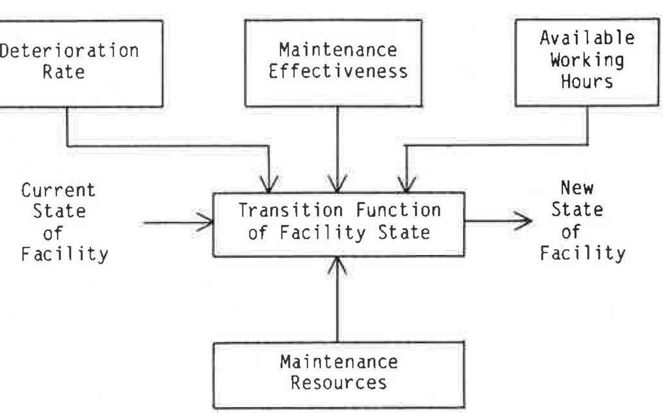

FIGURE 2 Elements of a maintenance model.

length of track (e.g., 10 m in the case of JNR) that exceed a predetermined critical value (e.g., 3 mm).

Each section of track has a length (ℓ_i) and a measure of importance, or weight (w_i), that generally reflects the volume and character of traffic over that section. Using these values, the objective of the problem may be formulated as follows:

$$\text{Minimize } T = \sum_{j=1}^{N} \sum_{i=1}^{n} w_i \ell_i P_{i,j+1} \quad (1)$$

where

- T = total weighted P-index over n sections and N periods,
- N = number of periods in the study period,
- n = total number of track sections,
- w_i = relative importance weight of section i,
- ℓ_i = track length of section i,
- $P_{i,j+1}$ = P-index of section i at the beginning of period j+1.

Changes in the P-index of section i are determined by usage of that section, its structural characteristics, climate, and maintenance resources allocated to it. In this case, the maintenance resource of interest is a tie-tamping machine that is used to level and align the track. To represent the dependence of the P-index on the decisions with respect to machine assignment, the transition function must be introduced. The transition function may be written as follows:

$$P_{i,j+1} = f_{ij}(P_{ij}, X_{ij} | d_{ij}, h_{ij}, e_{ij})$$

where

- f_{ij} = transition function of track state for section i, period j;
- X_{ij} = number of machines to be assigned to section i for period j;
- d_{ij} = deterioration rate of section i during period j;
- h_{ij} = available tamping hours in section i during period j; and
- e_{ij} = an index that specifies the determinants of the tamping effectiveness of section i during period j.

The first two arguments are independent variables and the last three are parameters. For ease of notation, this function will be written as

$$P_{i,j+1} = f_{ij}(P_{ij}, X_{ij}) \quad \text{for } i = 1,\ldots,n;\ j = 1,\ldots,N \quad (2)$$

Note, however, that in general the deterioration rate (d_{ij}) may be dependent on the current state (P_{ij}) in which case this notation is a slight oversimplification. Also, there may be several different types of resources (such as machines with different production rates). In this case, X_{ij} to X_{ijk} could be generalized and the resources of class k assigned to section i for period j could be denoted. This makes the problem larger (more variables) but does not alter the basic structure of the problem or the solution method. Thus Expression 1 may be rewritten as

$$\min T = \sum_{j=1}^{N} \sum_{i=1}^{n} w_i \ell_i f_{ij}(P_{ij}, X_{ij}) \quad (3)$$

This minimization is subject to certain constraints:

1. The total number of machines is limited in each period:

$$\sum_{i=1}^{n} X_{ij} \le M_j \quad \text{for } j = 1,\ldots,N \quad (4)$$

where M_j is the total number of machines available in period j.

2. The track state of each section must satisfy a safety standard specified for the section:

$$P_{i,j} \le UBP_i \quad \text{for } i = 1,\ldots,n;\ j = 1,\ldots,N \quad (5)$$

where UBP_i is the upper bound P-index prescribed for section i.

3. There may be upper or lower bounds, or both, on the number of machines to be assigned:

$$LBX_i \le X_{ij} \le UBX_i \quad \text{for } i = 1,\ldots,n;\ j = 1,\ldots,N \quad (6)$$

4. Track states at the beginning of period 1 (P_{i1}) are given.

Physically, the decision variables X_{ij} must be integers. However, they will be defined as real in this study. This means that some of the machines can be transferred in the middle of a period, or some of the machines are shared among two or more sections. Whether or not this assumption is totally appropriate depends on the system under study. Nevertheless, they are defined as real for three reasons: First, this assumption provides a lower bound on the integer solution. Second, the real solution may be converted to an integer solution relatively easily. Third, the solution provided by the mathematical technique is not always the final decision because there are still some elements that cannot be formulated correctly or cannot be formulated at all. Human judgment is still important in the final decision.

Expressions 3-6 define an optimization problem that is, in general, nonlinear because of nonlinearities in the transition functions (f_{ij}).

The formulation presented here implies the need for a model of a multiple-period decision process. However, it may be asked why all periods need to be considered simultaneously instead of optimizing the problem period by period. Because the decision involves only maintenance, it might appear that the smaller the value of the objective function up to a certain period, the easier it would be for the value of the objective function to be minimized for the remaining periods. To see whether this is true, the following hypothesis will be examined. Let the system state in period j be the vector with n elements: $P_j = (P_{1j}, P_{2j}, \ldots, P_{nj})$.

Hypothesis: If there are two system states in period j, P_j^1 and P_j^2, such that the total weighted P-index up to period j is smaller when the system state is in P_j^1, then the minimum total weighted P-index through period j+1 derived from the system state P_j^1 is smaller than that from the P_j^2 when all the possible decisions in period j+1 are considered.

If this hypothesis holds, the problem can be decomposed into N (the number of periods) problems of one period each. The optimum solution (the best policy) would be obtained by the sequence of the N short-run optimum decisions. Unfortunately, this does not hold in general. Consider the following small example, in which two sections are considered. The track length, weight, upper bound P-index, initial P-index, deterioration rate for each section, and available tamping hours for each period and each section are as given in Table 1. Suppose that the

TABLE 1 List of Data and Parameters

	Section 1	Section 2
Track length (km)	100.0	100.0
Weight	1.0	1.0
Upper bound P-index	30.0	40.0
Initial P-index	20.0	30.0
Deterioration rate (d_{ij}) (P-index points/period)		
Period 1	10.0	10.0
Period 2	10.0	10.0
Tamping hours (h_{ij})		
Period 1	100.0	100.0
Period 2	50.0	100.0

total number of available machines is two for each period. Assume that both track structure and topographical conditions are the same for both sections and that there are no lower or upper bounds on the number of machines to be assigned.

The transition functions f_{ij} (i = 1,2,; j = 1,2) are given according to the two parameters, deterioration rate and available tamping hours. They are shown in Figure 3. The result of computation is given in Table 2. This example shows that while the total weighted P-index by decision 3 up to period 1 has the minimum value (5,300), the minimum total weighted P-index (11,900) derived from the best decision in the second period is greater than the total weighted P-index from decision 1 in period 1 (11,500) and from decision 2 (11,300). An intuitive interpretation of this example is that because the number of available tamping hours of section 2 in period 2 is smaller, it is advantageous to assign machines to section 2 in period 1 rather than in period 2.

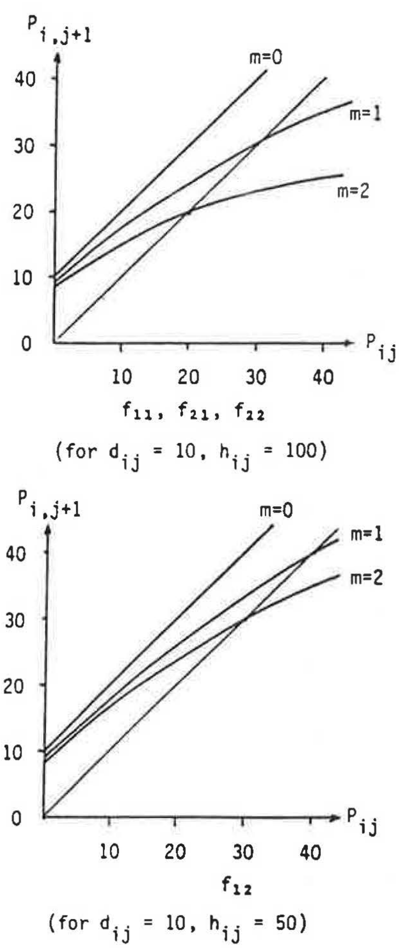

FIGURE 3 Example transition functions.

TABLE 2 Computed Results for Example Problem

Period 1 Decision[a]	Resulting Total Weighted P-Index	Period 2 Decision[a]	Resulting Total Weighted P-Index
1	6,000	1	Infeasible
		2	12,100
		3	11,500
2	5,400	1	12,000
		2	11,300
		3	Infeasible
3	5,300	1	11,900
		2	Infeasible
		3	Infeasible

[a] Decision 1 = assign both machines to Section 1, decision 2 = assign one machine to each section, and decision 3 = assign both machines to Section 2.

Hence the hypothesis is rejected. This simple example illustrates the need for a model of multiple periods, in which the dynamic aspects of the problem are represented properly.

SOLUTION OF THE PROBLEM

The model represented by Expressions 3-6 in the previous section is a nonlinear programming problem. In general, both the objective function and some of the constraints are nonlinear functions of the decision variables. In this section a procedure for obtaining solutions to the problem, using the generalized reduced gradient (GRG) method, is described.

The main idea of the GRG method is similar to that of the simplex method for linear programming. Using constraints, the vector of variables is partitioned into two subvectors: the vector of basic variables and the vector of nonbasic variables. The GRG method uses a gradient only in terms of the nonbasic vector, referred to as a reduced gradient, in order to improve the value of the objective function. For additional discussion of the GRG method, the interested reader is referred to Avriel (4).

To use the GRG method, it must be possible to compute partial derivatives of the objective function and all constraint expressions with respect to each of the decision variables. This can be done numerically, but, by using analytic expressions for the partial derivatives, both the speed and the accuracy of the algorithm can be improved. A discussion of the derivation of expressions for the required partial derivatives, using efficient recursive formulas, follows.

If $\{X_j\}$ is defined as the set of X_{ij} for $i = 1,\ldots,n$ (i.e., the allocations of machines to sections in period j), then $v_j(\{P_j\}, \{X_j\})$ can be defined as the value of the weighted P-index achieved in period j:

$$v_j(\{P_j\}, \{X_j\}) = \sum_{i=1}^{n} w_i \ell_i f_{ij}(P_{ij}, X_{ij}) \quad (7)$$

Note that $v_j(\{P_j\}, \{X_j\})$ is simply a subset of the terms in the objective function 1, corresponding to a specific period j.

Also

$$T_j = \sum_{k=j}^{n} v_k(\{P_k\}, \{X_k\}) \quad (8)$$

can be defined as the value of the objective function starting with period j and summed over the remaining periods, j, j+1,..., N. Note that a simple recursive equation can then be developed:

$$T_j = v(\{P_j\}, \{X_j\}) + T_{j+1} \quad (9)$$

The partial derivatives of the objective function (T) in Expression 3 with respect to X_{ij}'s can be computed as $\partial T/\partial X_{ij}$ (for i = 1,...,n, j = 1,...,N). Because a change in X_{ij} does not affect the value of T in stage 1 through state j-1, it is immediately apparent that the following is true:

$$\partial T/\partial X_{ij} = \partial T_j/\partial X_{ij}$$

Taking the partial derivative with respect to X_{ij} in Expression 9 yields

$$\partial T_j/X_{ij} = (\partial v_j/X_{ij}) + \sum_{\ell=1}^{n} (\partial T_{j+1}/P_{\ell,j+1})(\partial P_{\ell,j+1}/X_{ij}) \quad (10)$$

Substitution for $\partial P_{\ell,j+1}/\partial X_{ij}$ in Expression 10 yields

$$\partial T_i/\partial X_{ij} = (\partial v_j/\partial X_{ij}) + \sum_{\ell=1}^{n} (\partial T_{j+1}/\partial P_{\ell,j+1})(\partial f_{\ell j}/\partial X_{ij}) \quad (11)$$

Both $\partial v_j/\partial X_{ij}$ and $\partial f_{\ell j}/\partial X_{ij}$ can be computed; thus $\partial T_j/\partial X_{ij}$ can be computed if $\partial T_{j+1}/\partial P_{\ell,j+1}$ is known.

Expression 9 will again be used to compute this. Taking the partial derivatives with respect to $P_{\ell k}$ in Expression 9:

$$\partial T_j/\partial P_{\ell j} = (\partial v_j/\partial P_{\ell j}) + \sum_{k=1}^{n} (\partial T_{j+1}/\partial P_{k,j+1})(\partial f_{kj}/\partial P_{\ell j}) \quad (12)$$

Equation 12 shows the rule to compute $\partial T/\partial P$ at a certain stage by using $\partial T/\partial P$ at the next stage. The boundary condition at state N+1 is

$$\partial T_{N+1}/\partial P_{i,N+1} = 0 \qquad \text{for } i = 1,\ldots,n \quad (13)$$

Then, starting from stage N+1 and computing backwards by using Equations 11 and 12, $\partial T_j/\partial X_{ij}$ can be computed for all i's and j's.

In the case of the present problem, Equations 11 and 12 can be simplified by using the independence of state variables with respect to space; that is,

$$\partial P_{\ell,j+1}/\partial X_{ij} = \begin{cases} \partial f_{\ell j}/\partial X_{ij} & (\text{if } i = \ell) \\ 0 & (\text{if } i \neq \ell) \end{cases} \quad (14)$$

$$\partial P_{\ell,j+1}/\partial P_{ij} = \begin{cases} \partial f_{\ell j}/\partial P_{ij} & (\text{if } i = \ell) \\ 0 & (\text{if } i \neq \ell) \end{cases} \quad (15)$$

Using Expressions 14 and 15 in Equations 11 and 12, respectively,

$$\partial T_j/\partial X_{ij} = (\partial v_j/\partial X_{ij}) + [(\partial T_{j+1}/\partial P_{i,j+1})(\partial f_{ij}/\partial X_{ij})] \quad (16)$$

$$\partial T_j/\partial P_{ij} = (\partial v_j/\partial P_{ij}) + [(\partial T_{j+1}/\partial P_{i,j+1})(\partial f_{ij}/\partial P_{ij})] \quad (17)$$

Next, the partial derivatives of the constraints are computed with respect to X_{ij} (for i = 1,...,n; j = 1,...,N). The first set of constraints, given by Expression 4, can be differentiated easily. For the second set of constraints, given by Expression 5,

$$P_{i,j} = f_{ij}(P_{i1}, X_{i1}, X_{i2},\ldots,X_{i,j-1}) \quad (18)$$

Equation 18 shows that $P_{i,j}$ is a function of the initial state (P_{i1}) and the sequence of the decisions that have been made in the section up to period j-1.

Taking partial derivatives of P_{ij} in Equation 18 with respect to $X_{k\ell}$ (k = 1,...,n and ℓ = 1,...,N) yields

$$\partial P_{ij}/\partial X_{k\ell} = \begin{cases} \sum_{n_1=1}^{n} \sum_{n_2=1}^{n} \cdots \sum_{n_m=1}^{n} [(\partial P_{ij}/\partial P_{n_1,j-1})(\partial P_{n_1,j-1}/\partial P_{n_2,j-2})] \\ \ldots [(\partial P_{n_{m-1},\ell+2}/\partial P_{n_2,\ell+1})(\partial P_{n_m,\ell+1}/\partial X_{k\ell})] & (\text{if } j \geq \ell) \\ 0 & (\text{otherwise}) \end{cases} \quad (19)$$

Again using Equations 14 and 15,

$$\partial P_{ij}/\partial X_{k\ell} = \begin{cases} [(\partial P_{ij}/\partial P_{i,j-1})(\partial P_{i,j-1}/\partial P_{i,j-2})] \ldots [(\partial P_{i,\ell+2}/\partial P_{i,\ell+1}) \\ (\partial f_{i\ell}/\partial X_{i\ell})] & (\text{if } i = k \text{ and } j > \ell) \\ 0 & (\text{otherwise}) \end{cases} \quad (20)$$

Now all the partial derivatives required for the computation of the reduced gradient can be computed. Then, starting with an initial trial point and computing iteratively, the optimum solution can be obtained. The actual computations for the examples in this paper were done using the GRG2 package developed by Lasdon, Warren, and Ratnor (5).

ILLUSTRATIVE CALCULATIONS

Suppose that there are three track sections of interest with relative importance weights of 3.0, 2.0, and 1.0, respectively; these three sections may be assumed to be a main line, a sub-main line, and a local line, respectively. Considered will be the machine assignment planning for 1 year, divided into four periods; these four periods may be assumed to be four seasons (spring, summer, autumn, and winter).

Because of heavy snowfall during winter, available tamping hours are extremely limited in sections 2 and 3; on the other hand, tamping work is restricted during summer in section 1 because of the danger of rail buckling due to high temperature.

There are a total of 10 machines available per period. The management policy is to do tamping work no more than once within a period; that is, tamping distance executed in a certain section within a period is not to be more than the track length of the section.

Table 3 gives data for these three sections over the four periods. These data are based on values provided by JNR but do not necessarily correspond exactly with any specific sections of their system.

When X_{ij} machines are assigned to section i in period j, the distance tamped is computed as follows:

$$\ell_t = ch_{ij}X_{ij} \quad (21)$$

where

c = machine performance (distance/hour) and
h_{ij} = available tamping hours (hours/period).

The transition function $P_{i,j+1} = f_{ij}(P_{ij}, X_{ij})$ is given as

$$P_{i,j+1} = (ch_{ij}X_{ij}/\ell_i)\{-1 + [1 + 4a_i(P_{ij} + d_{ij}/2)]^{1/2}/2a_i\} + [1 - (ch_{ij}X_{ij}/\ell_i)]\,[P_{ij} + (d_{ij}/2)] + d_{ij}/2 \quad (22)$$

The derivation of this function is described in detail by Murakami (6).

TABLE 3 Data and Parameters for the Test Case

	Section		
	1	2	3
w_i	3.0	2.0	1.0
ℓ_i	225.3	241.4	217.3
a_i	0.05	0.025	0.015
UBP_i	35.0	37.0	39.0
$IP_i(=P_{i1})$	33.0	34.5	36.5
UBX_i		min(TM, ℓ_i/c h_{ij})	
LBX_i	0	0	0
Period			
1 d_{ij}	4.0	3.5	3.5
2	3.5	2.4	2.0
3	3.3	2.5	2.0
4	4.0	4.0	2.5
Avg	3.7	3.1	2.5
1 h_{ij}	50.0	60.0	60.0
2	40.0	80.0	100.0
3	50.0	60.0	80.0
4	40.0	30.0	20.0
Avg	45.0	57.5	65.0

Note: TM = total number of machines available per period, c = machine performance (= 0.32 km/hr), w_i = relative importance of section i, ℓ_i = track length of section i (km), a_i = tamping effect coefficient for section i, UBP_i = upper bound P-index for section i, $IP_i(=P_{i1})$ = initial P-index in section i, UBX_i/LBX_i = upper/lower bound number of machines to be assigned to section i, d_{ij} = deterioration rate of section i in period j, and h_{ij} = available tamping hours in section i during period j.

The optimal solution found for this problem is given in Table 4, and the best integer solution is given in Table 5. Note that the final P-index values for the integer and noninteger solutions are identical (to three significant figures). This indicates that the approximation involved in treating the problem with continuous variables is a reasonable one.

TABLE 4 Optimal Solution to Example Problem (machines assigned)

	Period			
Section	1	2	3	4
1	8.04	0.00	6.80	10.00
2	0.94	8.86	0.00	0.00
3	1.02	1.14	3.20	0.00

Note: Weighted average P-index after four periods = 29.1.

TABLE 5 Optimal Integer Solution to Example Problem (machines assigned)

	Period			
Section	1	2	3	4
1	7	0	7	10
2	1	9	0	0
3	2	1	3	0

Note: Weighted average P-index after four periods = 29.1.

As a further illustration of the effectiveness of the formulation described in this paper, the result of the optimization (as given in Table 4) has been compared with two alternative solutions:

- Assignment of a constant number of machines to each section for all periods (static solution) and
- Sequence of period-by-period optimal assignments (myopic solution).

Tables 6 and 7 show the results of these alternative solutions.

TABLE 6 Best Static Solution (machines assigned)

	Period			
Section	1	2	3	4
1	5	5	5	5
2	3	3	3	3
3	2	2	2	2

Note: Weighted average P-index after four periods = 33.0.

TABLE 7 Best Myopic Solution (machines assigned)

	Period			
Section	1	2	3	4
1	8.04	0.0	8.58	2.96
2	0.94	8.86	0.0	0.0
3	1.02	1.14	1.42	7.04

Noted: Weighted average P-index after four periods = 31.5.

The static solution yields a value of the weighted P-index at the end of the fourth period that is 10.8 percent worse than the optimal dynamic solution. The myopic solution is the same as the optimal solution in the first two periods and then differs over the last two periods. The myopic solution overallocates resources to section 1 in period 3, and then in period 4 it is forced to allocate most of the machines to section 3 in order to avoid violating the minimum standards in that section.

In contrast, the optimal solution looks at periods 3 and 4 together, recognizing that the machines can be used more effectively in section 1 in period 4 than in section 3. Thus it is seen that these two solutions are diverging, and, at the end of four periods, the myopic solution is about 6 percent worse than the optimal solution.

To illustrate the differences between the myopic period-by-period optimization and the dynamic optimization more completely, this test case was extended to 12 periods. The four-period seasonal cycle described in Table 3 was assumed to hold over two additional cycles, representing, in total, a 3-year planning horizon for the dynamic model.

Table 8 gives the machine assignments and resulting P-index values for the three sections over all 12 periods, allowing comparison of the myopic solution and the optimal solution. Note that the myopic solution follows a pattern of allocating just enough machine time to section 3 in each period to maintain the minimum required quality (maximum allowable P-index) with all remaining resources allocated to section 1 in the first, third, and fourth quarters of the year and to section 2 in the second quarter.

The optimal solution concentrates most maintenance in section 3 in two periods (2 and 7) and gen-

TABLE 8 Comparison of Myopic and Optimal Solutions

		Myopic Solution		Optimal Solution	
Period	Section	Machines Assigned	Resulting P-Index	Machines Assigned	Resulting P-Index
1	1	8.04	27.4	8.04	27.4
	2	0.94	37.0	0.94	37.0
	3	1.02	39.0	1.02	39.0
2	1	0	30.9	0	30.9
	2	8.86	26.0	3.25	34.5
	3	1.14	39.0	6.75	29.1
3	1	8.58	24.9	0	34.2
	2	0	28.5	10.0	26.6
	3	1.42	39.0	0	31.1
4	1	2.96	26.9	10.0	28.2
	2	0	32.5	0	30.6
	3	7.04	39.0	0	33.6
5	1	6.74	24.7	10.0	22.4
	2	0	36.0	0	34.1
	3	3.26	39.0	0	37.1
6	1	0	28.2	0	25.9
	2	8.86	25.5	9.38	23.8
	3	1.14	39.0	0.62	38.1
7	1	8.58	23.3	1.56	27.9
	2	0	28.0	0	26.3
	3	1.42	39.0	8.44	28.7
8	1	2.96	25.4	10.0	24.2
	2	0	32.0	0	30.3
	3	7.04	39.0	0	31.2
9	1	6.74	23.7	10.0	20.2
	2	0	34.5	0	33.8
	3	3.26	39.0	0	34.7
10	1	0	27.2	0	23.7
	2	8.86	24.7	9.38	23.6
	3	1.14	39.0	0.62	35.7
11	1	8.58	22.7	9.0	20.1
	2	0	27.2	0	26.1
	3	1.42	39.0	1.0	36.5
12	1	2.96	25.0	10.0	19.1
	2	0	31.2	0	30.1
	3	7.04	39.0	0	39.0

Note: Final average weighted P-index is 29.4 for the myopic solution and 26.1 for the optimal solution.

erally shows the pattern of using most resources in a single section in each period, rotating among the sections and keeping them all below the maximum allowable P-index. The weighted P-index achieved after 12 periods is 26.7, some 11 percent better than the myopic solution.

CONCLUSIONS

The formulation of a dynamic model for allocating maintenance resources across several facilities (or facility sections) over multiple time periods has been described. The basic structure of the model includes deterioration rates of the facilities over time, capability to represent variable effectiveness of maintenance resources in various sections over time, and variable amounts of total resources available over time.

The model has been formulated as a nonlinear optimization problem, and a solution method using the generalized reduced gradient approach has been described. This solution method has modest computation requirements and can be implemented using either commercially available software packages or custom software.

The formulation described here is based on using a single composite measure of facility condition (or state). The PSI rating for pavements and the P-index of track surface condition are examples of such measures. However, it must be recognized that single composite measures do not always represent all elements of a maintenance problem. They are simply an overall guide to relative facility condition.

In some applications, the state variables (facility condition measures) may not be independent from one section to another as assumed in the example in the previous section. The model can handle this complexity quite easily--the only change is that Equations 11, 12, and 19 are used in place of 16, 17, and 20 to compute the required derivatives.

An illustrative example based on data from Japanese National Railways has shown how the solution obtained from the optimization model compares to alternative, simpler, solutions: a static solution of doing the same thing each period and a myopic solution of optimizing one period at a time. The differences shown, even in this small example, emphasize the value of considering the dynamic elements of the problem.

The model described here is applicable to the analysis of maintenance planning problems in a variety of situations, including highway and waterway maintenance as well as railroad applications. For example, this model would be an effective complement to the Highway Condition Projection Model (HCPM) described by Hartgen (7) and used in New York State. That model is descriptive in nature, a "what-if" tool that predicts the effects of analyst-specified maintenance strategies. The model described here is an optimization model useful in designing those strategies.

In summary, this research provides an important new tool for the use of maintenance planners with the potential of helping make more efficient use of maintenance resources in many different applications.

REFERENCES

1. Final Report on the Federal Highway Cost Association Study. FHWA, U.S. Department of Transportation, 1982.
2. Railroad Facts, 1983 Edition. Association of American Railroads, Washington, D.C., 1983.
3. Interim Guide for Design of Pavement Structures. AASHTO, Washington, D.C., 1981.
4. M. Avriel. Nonlinear Programming: Analysis and Methods. Prentice-Hall, Englewood Cliffs, N.J., 1976.
5. L. Lasdon, A. Warren, and M. Ratnor. GRG2 User's Guide. Department of General Business, University of Texas, Austin, 1980.
6. K. Murakami. A Multiple Period Optimization Model for Scheduling Maintenance of Transportation Facilities. M.S. Thesis. School of Civil and Environmental Engineering, Cornell University, Ithaca, N.Y., 1984.
7. D. Hartgen. Long-Term Projection of Highway System Condition. In Transportation Research Record 940, TRB, National Research Council, Washington, D.C., 1984, pp. 8-15.

Publication of this paper sponsored by Committee on Railway Maintenance.