Optimizing highway maintenance operations: dynamic considerations

Saeideh Fallah-Fini, a* Hazhir Rahmandad, b Konstantinos Triantis b and Jesus M. de la Garza c

Abstract

Effective highway maintenance depends on several activities, including the understanding of current and the prediction of future pavement conditions and deciding how to best allocate limited resources for maintenance operations. In this paper, a dynamic micro-level simulation model of highway deterioration and renewal processes is presented. This model is calibrated with data from eight road sections in Virginia and is coupled with an optimization module that optimizes maintenance operations. The analysis offers alternative priority setting schemes that improve current maintenance practices at the project and network levels. This approach provides a blueprint for designing optimal highway maintenance practices.

Introduction

Gradual deterioration of the U.S. road infrastructure in the past two decades (ASCE, 2009a), along with major budgetary restrictions for making significant improvements in road conditions (ASCE, 2009b), have raised the importance of improving highway maintenance practices. In fact, the Federal Highway Administration (FHWA) has endorsed “asset management” as the future approach for road maintenance in all state Departments of Transportation (DOTs) (JLARC, 2002). Asset management calls for the utilization of engineering, management, and economics principles to help state DOTs with allocating limited resources to preserving, operating, and managing the nation’s road infrastructure (Ozbek, 2007).

This focus, over the past two decades, has led to the development of pavement management systems (PMSs) (Gendreau and Soriano, 1998) that draw on systematic methods to best use available budgets and to design cost-effective maintenance policies. One of the key components of a PMS is planning methods to decide which maintenance operations should be performed when considering the current and predicted condition of the pavement (Gendreau and Soriano, 1998). The current research contributes to this component of a PMS by using the system dynamics approach to find the optimal policy for allocating the maintenance budget given a set of environmental and operational conditions.

* Correspondence to: Saeideh Fallah-Fini. E-mail: fallah@vt.edu

Received 10 February 2010; Accepted 25 June 2010

Copyright © 2010 John Wiley & Sons, Ltd
Figure 1 specifies the major steps in our research. First, we develop a simulation model of the highway deterioration process at a micro (road section) level using the principles embodied in mechanistic-empirical models (Huang, 2004), which are among the best available physics-based models for predicting road deterioration in pavement management. Second, we use empirical data (pavement condition and traffic for approximately 17 miles of Virginia's interstate highway during the years 2002–2007) to estimate the unknown parameters of this model. Finally, we couple the calibrated pavement deterioration model with the highway maintenance and renewal decision-making process to find the best policy for allocating the maintenance budget to different maintenance operations. Reliable policy recommendations require optimization to be performed in robust and calibrated models. Thus, this paper couples system dynamics modeling with calibration and optimization and builds a concrete framework for designing maintenance policies across different settings.

The remainder of this paper is organized as follows. The approaches that exist in the literature for the planning of maintenance operations are described in the next section. The third section describes the methods used for modeling the dynamics of road deterioration and maintenance as well as the calibration and optimization of the model. The results and insights are discussed in the fourth section. Conclusions and future directions are provided in the fifth section.

Approaches for planning of road maintenance operations

In the past two decades, several approaches such as expert systems, analytical models and system dynamics have been utilized to formulate highway maintenance planning and scheduling. With the advances in expert knowledge elicitation and analysis methods, “expert systems” have become popular for pavement management in both highway and airport networks (Ismail et al., 2009). One of the advantages of expert systems is their ability to involve expert knowledge and subjective human reasoning (Chang Albitres et al., 2005) that typically are complicated to incorporate into mathematical models. However, lack of agreement among experts in the field and inadequate approaches for representing domain knowledge remain as major shortcomings (Ismail et al., 2009).

The second approach for highway maintenance scheduling and budget planning is analytical modeling. These approaches formulate the decision-making problem as a model where the objective function (e.g. minimizing overall costs) and the constraints (e.g. minimum acceptable condition) are represented by mathematical expressions of the decision variables (Gendreau and Soriano, 1998). These mathematical modeling techniques are either deterministic (Lytton, 1985; Feighan et al., 1987; Fwa et al., 1988, 2000; Smadi, 1994; Wang et al., 2003) or stochastic (Butt et al., 1994; Gao et al., 2007; Gao and Zhang, 2008). They can often handle problems with major detail complexity, yet they also suffer from (i) extensive computational requirements; (ii) lack of prior knowledge of the data generating process that is required for stochastic modeling of the pavement deterioration process; and (iii) the gap between theory and practice that has led to the challenge of interpreting complex analytical models by maintenance managers (Dekker, 1996; Gao and Zhang, 2008).

The third approach is the application of system dynamics (SD). Road deterioration and maintenance is a dynamic phenomenon embedded in a nonlinear feedback system
Fig. 1. The three main steps of this research study toward optimizing highway maintenance policies.
where the uncontrollable factors of the environment (e.g. climate condition) and operational conditions (e.g. traffic load) as well as controllable factors (e.g. maintenance policy) affect the road condition. SD is an appropriate modeling approach for this context because it captures the dynamics of road conditions and accounts for feedback loops that determine the physical road deterioration processes. There are a few applications of SD to transportation systems (Ogunlana et al., 1998, 2003; Ibbs and Liu, 2005; Lee and Pena-Mora, 2005; Lee et al., 2005; Thompson and Bank, 2010) and highway maintenance policy (Chasey, 1995; Chasey et al., 1997, 2002; de la Garza et al., 1998; Kim, 1998; Bjornsson et al., 2000; de la Garza and Krueger, 2007). For example, the impact of deferred maintenance on the interstate's level of availability (i.e. in terms of capacity and allowable traffic volume) and on the level of operation (i.e. the road physical condition) has been modeled by a number of researchers (Chasey, 1995; Chasey et al., 1997, 2002; de la Garza et al., 1998). In addition, at the macro level, the effects of deteriorated roads on user benefits (i.e. vehicle operating cost and travel time) and non-user benefits (e.g. increase in business opportunities) have been considered.

All of the prior SD highway maintenance studies are at the network level, meaning that the unit of analysis is the overall highway network under the control of the road authority. In the real world, the whole highway network is divided into several road sections and maintenance treatments are defined for each one of these sections. Thus, the road section is the smallest unit for which the maintenance decisions are made. Therefore, network-level models are sometimes too aggregated to provide practical decision support at the road section level. In fact, maintenance managers would like decision support both at the road section and the overall network levels.

In addition, in previous SD highway maintenance studies, model parameters such as deterioration rates are defined either based on synthetic data or using the outputs of other statistical analyses. Therefore, the potential to estimate model parameters using original data has not been realized and thus ensuring model reliability has been a challenge. Furthermore, previous studies have relied on sensitivity analysis and trial and error for policy analysis and have not used an integrated optimization method to find the best maintenance policies. This paper advances the application of SD to highway maintenance by addressing these shortcomings.

**Methods**

At an aggregate level, deterioration and maintenance dynamics can be summarized in two major feedback loops (Figure 2). The pavement condition is deteriorated by traffic load, climate condition, and other deterioration factors. The balancing loop B1 (Maintenance Fix) describes how maintenance operations performed by road authorities attempt to bring the road condition towards desired conditions by reducing the area under distress. On the other hand, the reinforcing loop R1 (Accelerated Deterioration) illustrates the effect of the budget shortfall on the delayed maintenance and further deterioration of pavement conditions. Note that although the physics of road deterioration and maintenance are complicated, the corresponding feedback structure is relatively simple. The real challenge lies in building a reliable model that quantifies the road deterioration process and corresponds to empirical findings. This is the challenge we tackle in the following sections.
Modeling the dynamics of road deterioration

Given that the goal of this study is to use the model for policy analysis through optimization, a modeling framework that is robust to extreme conditions and one that is understood by policy makers is required. To this end, the pavement engineering literature was explored to represent the physics of pavement deterioration (Huang, 2004). Within the SD framework (Forrester, 1971; Sterman, 2000), these physically based dynamics were studied in conjunction with macro-level maintenance operations. This combination allows for building a simulation model that is grounded in the physics of road operations that can be validated (Fallah-Fini and Triantis, 2009) and which considers social and managerial factors. We seek a broad model boundary that is robust to extreme input levels and that allows for reliable policy analysis (Sterman, 2000).

The pavement deterioration module of the SD model captures distress generation and propagation on road sections. It is a critical component of the model, since it forms the main input to the maintenance decision-making process and can significantly affect the effectiveness of the recommended maintenance policies. This module relies on mechanistic–empirical (ME) models (Huang, 2004) to capture the road deterioration process. ME models use mechanical analysis to design the pavement structure. Pavement structure determines the “allowable number of load cycles”. This parameter represents the number of load cycles a pavement can bear before a specific percentage of the pavement area experiences distress.
The damage ratio, which is the ratio of the “actual” to the “allowable” number of load cycles, shows the damage corresponding to a specific condition (e.g. traffic load). Damage to the pavement accumulates as the actual number of load cycles increases. In practice, it is common to consider that when the ratio reaches 1.0, 20 percent of the lane area experiences fatigue cracking (Priest and Timm, 2006). In this study, it is assumed that the damage that gradually appears on a road section is proportional to the damage ratio.

The allowable number of load cycles is specified for each type of distress. For flexible pavements, such as those in highways, the main load-related distresses include fatigue cracking and rutting. Fatigue cracking is a series of interconnected cracks caused by fatigue failure of the pavement surface due to repeated traffic loading and is considered as the major structural distress on the road surface (Priest and Timm, 2006). In this study we therefore focus on fatigue cracking.

The actual number of load cycles is calculated by considering the availability of traffic and vehicle classification data for each road section; see Fallah-Fini et al. (2009) for details about calculating the actual number of load cycles. The allowable number of load cycles is calculated using the data related to pavement structure at each road section. These data require detailed knowledge of road construction and material, not typically available for older roads. Thus, in this SD model, the “allowable” number of load cycles is an unknown parameter that is estimated through calibration.

The Virginia Department of Transportation (VDOT) specifies fatigue cracking levels by two parameters: severity and density of the cracks (VDOT, 2002). Severity, which represents depth and openness of a crack, is defined at three levels: Not-Severe (NS), Severe (S), and Very-Severe (VS). Density of the cracks represents the percentage of the right lane area that is covered with cracks. The three density levels that are used in practice are Rare (less than 10 percent of the right lane area is filled with the cracks), Occasional (between 10 and 50 percent), and Frequent (more than 50 percent).

When a road section is evaluated, the density levels corresponding to all three fatigue cracking severity levels (NS, S, and VS) are defined. For example, a road section may contain 5 percent NS cracks, 10 percent S cracks, and 5 percent VS cracks. The overall condition of a road section at any point in time is defined by: (1) choosing the severity level with the maximum density (i.e. S in the example); (2) adding up the density of the three severity levels to determine the overall density level (i.e. 20 percent in the example). Thus the overall condition of the example road section is reported as “Severe & Occasional”. In addition, if a road section is in a perfect condition, it is represented as “Not-Severe & None”, meaning that no distress has been observed in the road section. These described steps are referred to as the “road condition evaluation procedure” hereafter. This procedure is performed annually for all road sections across a highway network. Table 1 shows the possible combinations of severity and density levels in a three-by-three matrix.

In this study the damage ratio is used to model the generation of fatigue cracking through the life cycle of a road section. It is assumed that a crack is in an NS condition when it is initially generated. Deferred maintenance operations due to lack of maintenance budget and suboptimal maintenance policies with repeated traffic loads cause the continuous transition of lane miles under distress from NS to S and from S to VS conditions, respectively. Thus three stocks are defined for the three severity levels of fatigue cracking (see Figure 3).
There is a delay between applying maintenance treatments on each road section and recurrence of distress on that road section. Thus, the road sections that are maintained are reassigned to the “Not-Severe & None” condition and remain in that state for some time. The “maintained lane miles” stock is defined to accumulate the maintained lane miles that are depleted from the three stocks of distresses. The outflow of the “maintained lane miles” stock releases the maintained lane miles into the deterioration process.

Time constants representing the speed of transfer across different stocks depend largely on the structure of the road sections and environmental conditions. Thus, these parameters are estimated in the calibration process. Figure 3 shows the simplified structure of the pavement deterioration module, its main variables, inputs and outputs. The four parameters that are computed by the calibration process are highlighted in bold and italic type and the exogenous inputs coming from exogenous time series data are underlined and highlighted in bold.

**Model calibration**

In this section a model-based calibration method for parameter estimation (Oliva, 2003) is used to find appropriate values for unknown parameters. To analyze road maintenance operations at a project level, the units under analysis are defined as road sections. By modeling multiple road sections, analysis of network-level metrics is also possible. The highway network in this study is represented by eight road sections that lie in a county in the state of Virginia.

VDOT uses the load-related distress (LDR) index to reflect the condition of the road with respect to distresses that are mainly caused by traffic load, such as fatigue cracking and rutting. LDR values range between 0 and 100, where an LDR of 100 represents a road section in perfect condition. To obtain the LDR for each road section, one subtracts from 100 the deduct value corresponding to each one of the nine possible road conditions presented in Table 1. Table 2 shows the standard deduct values for fatigue cracking developed by VDOT’s Asset Management Division (Chowdhury, 2007). Based on Table 2, the LDR value of a road section that is in “Severe & Rare” condition is reported as $87 = 100 - 13$.

Table 3 summarizes the road condition data for the eight road sections in the highway network under analysis after some data cleaning. These data points are used for the calibration process. The cells highlighted in bold face show the years in which maintenance operations were performed. As the data show, when a maintenance operation

<table>
<thead>
<tr>
<th>Severity level</th>
<th>Density level</th>
<th>Rare (R)</th>
<th>Occasional (O)</th>
<th>Frequent (F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-Severe (NS)</td>
<td>NS-R</td>
<td>NS-O</td>
<td>NS-F</td>
<td></td>
</tr>
<tr>
<td>Severe (S)</td>
<td>S-R</td>
<td>S-O</td>
<td>S-F</td>
<td></td>
</tr>
<tr>
<td>Very-Severe (VS)</td>
<td>VS-R</td>
<td>VS-O</td>
<td>VS-F</td>
<td></td>
</tr>
</tbody>
</table>

There is a delay between applying maintenance treatments on each road section and recurrence of distress on that road section. Thus, the road sections that are maintained are reassigned to the “Not-Severe & None” condition and remain in that state for some time. The “maintained lane miles” stock is defined to accumulate the maintained lane miles that are depleted from the three stocks of distresses. The outflow of the “maintained lane miles” stock releases the maintained lane miles into the deterioration process.

Time constants representing the speed of transfer across different stocks depend largely on the structure of the road sections and environmental conditions. Thus, these parameters are estimated in the calibration process. Figure 3 shows the simplified structure of the pavement deterioration module, its main variables, inputs and outputs. The four parameters that are computed by the calibration process are highlighted in bold and italic type and the exogenous inputs coming from exogenous time series data are underlined and highlighted in bold.

**Model calibration**

In this section a model-based calibration method for parameter estimation (Oliva, 2003) is used to find appropriate values for unknown parameters. To analyze road maintenance operations at a project level, the units under analysis are defined as road sections. By modeling multiple road sections, analysis of network-level metrics is also possible. The highway network in this study is represented by eight road sections that lie in a county in the state of Virginia.

VDOT uses the load-related distress (LDR) index to reflect the condition of the road with respect to distresses that are mainly caused by traffic load, such as fatigue cracking and rutting. LDR values range between 0 and 100, where an LDR of 100 represents a road section in perfect condition. To obtain the LDR for each road section, one subtracts from 100 the deduct value corresponding to each one of the nine possible road conditions presented in Table 1. Table 2 shows the standard deduct values for fatigue cracking developed by VDOT’s Asset Management Division (Chowdhury, 2007). Based on Table 2, the LDR value of a road section that is in “Severe & Rare” condition is reported as $87 = 100 - 13$.

Table 3 summarizes the road condition data for the eight road sections in the highway network under analysis after some data cleaning. These data points are used for the calibration process. The cells highlighted in bold face show the years in which maintenance operations were performed. As the data show, when a maintenance operation
Fig. 3. Simplified structure of the pavement deterioration module
is performed on a road section, the condition of the road goes back to Not-Severe & None state in the next year. Thus, when a maintenance operation is performed in the calibration, the three stocks of distress are fully depleted. Maintenance happens exogenously in the calibration runs so as to be consistent with the data.

The objective of the calibration process is to determine the values of the four unknown parameters presented in Table 4 so that the simulated road sections follow the data provided in Table 3. The payoff function that is minimized in the calibration is the weighted sum of eight error terms for different road sections. These error terms are zero if the road section condition in the model equals what is observed in the data and one otherwise. Error terms are weighted proportional to the inverse of the standard deviation of the error terms, which leads to maximum likelihood estimates for the estimated parameters (Greene, 2002). Table 4 shows the calibration results.

The calibration results are in a reasonable range considering the road condition data in Table 3. For example, Table 3 shows that most of the road sections remain in Not-Severe & None condition for 2 years before recurrence of distresses. As can be seen, the estimated value for the corresponding parameter is 23.7 months. Figure 4, as an example, compares the data and model estimates for LDR in road sections 4 and 6. Figure 5 shows the behavior of the LDR variable at the network level and its comparison with the network-level LDR variable obtained from the data. The network-level LDR was obtained as the weighted average of the LDR of the road sections, where the lengths of the road sections were used as the weight factor.

Note that data points only record the road condition at the beginning of each year, whereas the model behavior is continuous. For example, in Figure 5 data (dots) at months 48 and 60 show that network-level LDR has changed from 93 in year 2006 (month 48) to 88 in year 2007 (month 60), while the model (line) shows that the network-level road condition has continuously deteriorated through the year 2006 to the start of the year 2007. Also note that LDR takes different constant values over different periods of time because of the calculation method used by VDOT (see Table 2). As a result, LDR values change discretely even though their driving variables are continuous.

**Modeling and optimization of the maintenance budget allocation**

In this section the road deterioration module is coupled with a “maintenance budget allocation” module to capture the dynamics of the budget allocation for various maintenance operations for each road section. The dynamics that are described in this section mostly relate to loop B1 (see Figure 2), where deteriorations generated in the

<table>
<thead>
<tr>
<th>Table 2. Standard deduct values for fatigue cracking developed by VDOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rare (R)</td>
</tr>
<tr>
<td>Not-Severe (NS)</td>
</tr>
<tr>
<td>Severe (S)</td>
</tr>
<tr>
<td>Very-Severe (VS)</td>
</tr>
</tbody>
</table>
Table 3. Road condition data over the years 2002–2007 for road sections in the network under analysis

<table>
<thead>
<tr>
<th>Road section ID</th>
<th>Length (mile)</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Severity</td>
<td>Density</td>
<td>LDR</td>
<td>Severity</td>
<td>Density</td>
<td>LDR</td>
</tr>
<tr>
<td>1</td>
<td>0.85</td>
<td>S</td>
<td>R</td>
<td>87</td>
<td>S</td>
<td>R</td>
<td>87</td>
</tr>
<tr>
<td>2</td>
<td>3.49</td>
<td>S</td>
<td>R</td>
<td>87</td>
<td>S</td>
<td>R</td>
<td>87</td>
</tr>
<tr>
<td>3</td>
<td>0.77</td>
<td>NS</td>
<td>None</td>
<td>100</td>
<td>NS</td>
<td>None</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>2.12</td>
<td>S</td>
<td>O</td>
<td>70</td>
<td>NS</td>
<td>None</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>2.16</td>
<td>S</td>
<td>R</td>
<td>87</td>
<td>S</td>
<td>O</td>
<td>70</td>
</tr>
<tr>
<td>6</td>
<td>0.75</td>
<td>S</td>
<td>O</td>
<td>70</td>
<td>S</td>
<td>O</td>
<td>70</td>
</tr>
<tr>
<td>7</td>
<td>1.07</td>
<td>S</td>
<td>R</td>
<td>87</td>
<td>S</td>
<td>O</td>
<td>70</td>
</tr>
<tr>
<td>8</td>
<td>5.65</td>
<td>S</td>
<td>R</td>
<td>87</td>
<td>NS</td>
<td>None</td>
<td>100</td>
</tr>
</tbody>
</table>

S. Fallah-Fini et al.: Optimizing Highway Maintenance Operations
Table 4. Estimated values for the parameters of interest obtained from the calibration process

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimated value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allowable number of load cycles</td>
<td>1e+8 load cycles</td>
</tr>
<tr>
<td>Time delay for transition from Not-Severe cracks to Severe cracks</td>
<td>10.9 months</td>
</tr>
<tr>
<td>Time delay for transition from Severe cracks to Very-Severe cracks</td>
<td>20.1 months</td>
</tr>
<tr>
<td>Time delay between finishing maintenance operations and reappearance of cracks</td>
<td>23.7 months</td>
</tr>
</tbody>
</table>

Road sections (as discussed and modeled above: “Modeling the dynamics of road deterioration”) are evaluated to define the required maintenance operations for each road section. The limited available budget is then allocated to the required maintenance operations based on the policy factors that are discussed and optimized using the optimization module. Thus the optimization module finds an optimum balance between performing maintenance operations (loop B1) to avoid accelerated deterioration (loop R1) and to save on the costs of maintenance. In what follows, the modeling and optimization of the budget allocation processes are discussed.

The required maintenance activities (preventive, corrective, and restorative maintenance) are defined based on the types of distresses on a road section, as well as the distress's severity and density levels. Table 5 shows the decision matrix used by VDOT (Chowdhury, 2007) to determine the maintenance activity suitable for a road section in the presence of fatigue cracking. For example, the maintenance strategy for the “Severe & Occasional” condition is preventive maintenance. Preventive maintenance (PM) refers to the treatments that are performed to reduce the rate of deterioration and preserve the existing pavement integrity. Corrective maintenance (CM) refers to the treatments that maintain the characteristics and structural integrity of an existing pavement for continued serviceability. Restorative maintenance (RM) refers to new surface layers that restore the pavement structure to a level similar to its original condition (Smith and Nazarian, 1992; Kim, 1998). Rehabilitation and reconstruction activities are not considered in this study. Costs for different maintenance activities, in dollars per lane mile of pavement, were estimated by investigating the literature (Mahoney et al., 2010) and consulting with practitioners. When an area within a road section needs to be maintained, the entire road section is maintained to restore the pavement to the “Not-Severe & None” condition.

Highway maintenance budgets are limited, and thus not all road sections will be maintained every year. What follows represents the real-world allocation of a limited network-level maintenance budget among the road sections. First, Table 1 and Table 2 are used to define the LDR of each road section. The road section’s LDR is then compared with an “LDR threshold” set by road authorities so that only those road sections with LDRs below the threshold are eligible to receive maintenance that year. For example, with an LDR threshold of 80, a road section that has an LDR of 90 is not eligible for maintenance. Next, Table 5 is used to determine the appropriate maintenance operations (i.e. PM, CM, and RM) for road sections that are eligible for maintenance. As a result, the total required budgets for PM, CM, and RM in the highway network are determined. The next step is to allocate the limited network-level budget among competing maintenance operations, i.e. PM, CM, and RM, based on their priorities.
Fig. 4. Comparing the LDR index obtained from the real data and the model after calibration for the road sections 4 and 6.
The `allocate available` function of VENSIM was used to model the allocation process. This function takes as its argument the network-level available budget, the total budget required by the PM, CM, and RM operations, and the priority profiles corresponding to each of these maintenance operations. The priority profile is represented by a normal distribution curve with a given mean (priority) and standard deviation (width), where the area under the curve represents the demand (required budget to perform the specific type of maintenance). The priority and width defined for the three different maintenance operations determine the budget allocation policy.

The following illustration shows how for given priority and width levels the `allocate available` function allocates the budget among multiple demands. Consider a case in which a limited maintenance budget of $60,000 needs to be allocated between PM and CM, where PM requires $40,000 and CM requires $30,000. Two scenarios are depicted in Figure 6: (a) PM and CM have priorities of 7 and 4 and widths of 0.5 and 1, respectively; (b) priorities and widths are 6 and 1 for PM and 4 and 2 for CM. Once the normal curves with these profiles are portrayed graphically, the allocation function works by distributing the resource (budget in this case) to the areas under the curves from right to left, until all the resource (budget) is allocated or all demand curves...
are fully satisfied. In these graphs the resulting budget allocations are shown by shaded areas under the curves for PM and CM, and the exact values are reported in Figure 6.

In cases similar to Figure 6(a), when the allocated budget for a maintenance operation (i.e. allocated budget for CM in this example) is not sufficient for fixing all road sections in need of that operation, the road sections that are in a worse condition (i.e. have lower LDRs) have a higher priority for receiving the limited available budget.

There are several decision parameters in the procedure described above that can potentially affect the outcome of budget allocation and the final condition of the road sections. The base case values of these parameters are used to represent current practice, and an optimization module is used to find improvements to the base case policy that can help road authorities make the best use of limited available resources.

The first set of parameters is related to the priority profiles of PM, CM, and RM. As shown in Figure 6, assigning different values to the mean (priority) and standard deviation (width) of the demand curves can lead to different scenarios for allocating the limited available budget among the maintenance operations and the corresponding road sections. Based on discussion with maintenance managers, CM usually receives a
higher priority. Finding the optimal set of priorities for PM, CM, and RM and their sensitivity with respect to the available network-level budget is analyzed below.

The second set of parameters is related to the density thresholds (used in Table 1) for defining Rare vs. Occasional (i.e. when 10 percent of the road section is covered with distress in practice) and Occasional vs. Frequent (i.e. when 50 percent of the road section is covered with distress in practice). By changing these thresholds, the conditions under which a road section goes from one density category to the other, and thus the required maintenance activities, are changed. However, one cannot change these thresholds arbitrarily since very high density threshold values may make it infeasible to perform the necessary maintenance operations. Therefore, as will be discussed in the next section, some feasible ranges for these parameters are set based on expert input and the optimum values are determined through optimization.

The third set of parameters that can be evaluated through optimization is the type of maintenance operation recommended for each cell in Table 5. Again, not all alternative treatments are feasible for each cell. However, when the recommended treatment is “Do Nothing”, one can always apply PM instead and test whether this can improve the overall performance of the road network.

Finally, the fourth parameter is the LDR threshold. In practice, this parameter is defined based on the availability of the maintenance budget. Road authorities that are facing a shortage in their budget choose lower LDR thresholds so as to be able to assign the limited budget to those road sections that are in critical need of maintenance, i.e. have very low LDRs. After setting a feasible range for this parameter based on expert input, optimization is used to find its best value.

In general, an “optimized” policy is compared with a “base case” scenario. The decision rules that are currently used by road authorities are utilized to construct the base case scenario (see Table 6). It is postulated that by using the calibrated model as discussed above (“Model calibration”) and an optimization method that varies the parameters discussed in this section, an optimal use of the maintenance budget is possible. Specifically, the optimization module of VENSIM is used in this study to find the optimum values of the parameters such that a desired utility function is minimized. The utility function in this paper is defined as the weighted sum of the area of the highway network which is under distress with different severity levels. The optimal maintenance budget allocations are defined such that the sum of the utility function over the analysis horizon is minimized.

Another complicating factor concerns the initial conditions of the road sections. The historical initial conditions that were observed for this network may lead the optimization algorithm to a budget allocation policy that is not generally reasonable. To account

<table>
<thead>
<tr>
<th>Parameters considered in optimization</th>
<th>Values in the base case scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priorities of the PM, CM, and RM</td>
<td>First priority: CM; second priority: PM; third priority: RM</td>
</tr>
<tr>
<td>Density threshold used in Table 1</td>
<td>10% and 50% for Rare and Occasional, respectively</td>
</tr>
<tr>
<td>Variations of Table 5</td>
<td>Table 5 is used as it is in the base case</td>
</tr>
<tr>
<td>LDR threshold</td>
<td>85</td>
</tr>
</tbody>
</table>
for this possibility, the model is subscripted for 50 replications of the same road network, which differ only in their initial values of stocks. To be realistic, the initial values for the stocks are randomly drawn from the distribution of road conditions over the calibrated base case simulations. For the optimization, the value of the utility function is summed over all 50 road network configurations, rather than focusing on a single network configuration with only one set of initial conditions. Therefore, the resulting solution is assumed to be robust to the initial road conditions of the three stocks. The optimization results and corresponding discussion are presented next.

**Optimization results and discussion**

State DOTs are interested in identifying the best maintenance resource allocation alternatives given different budgetary scenarios. Therefore, in order to find these alternatives, four budget scenarios are considered, namely, “Low”, “Most likely”, “High”, and “Extreme”. The values associated with these scenarios (see Table 7) were found by investigating the literature (Mahoney et al., 2010) and consulting with practitioners. For each budget scenario, the optimization results are presented in Table 7.

A few interesting observations emerge. First, RM is rarely needed for any of these scenarios since the optimally allocated budget is sufficient to avoid deterioration of the road sections to “Very-Severe & Frequent” where road sections would need RM. As a result, the RM profiles (with their respective priorities and widths) are not reported in Table 7.

Second, overall, PM gets a higher priority when compared with CM in the optimized solutions. This result is robust when considering the four different budget levels. In essence, by applying PM, the road authorities can effectively decrease the need for future CM while spending less overall. However, the overlap in the PM and CM profiles increases as the total budget goes down. This shows the need for sharing the budget between PM and CM, rather than satisfying PM first and then allocating leftover resources to CM. This overlap is justified when one considers that with low maintenance budgets not all PM work can be completed, which logically leads to an increase in the required CM work in subsequent years.

It is interesting to note that the current resource allocation practice favors CM over PM, which is not optimal according to this analysis. In practice, however, the preference of CM over PM likely reflects a short-term perspective vis-à-vis a long-term one. That is, road maintenance engineers struggle with the difficult notion of allowing current road sections requiring CM to further deteriorate, even if they know that conducting PM first is better over the long run.

The impact of the distress density thresholds for defining Rare vs. Occasional (Rare Upper Bound (UB)) and Occasional vs. Frequent (Occasional Upper Bound (UB)) is more subtle, even though these thresholds have a noticeable impact on the utility function because they influence how different road sections are categorized. Whether a road section is seen as highly deteriorated and in need of maintenance or not depends on these thresholds. For example, consider road sections A and B in which “Very-Severe” cracks have the highest density, with 7 percent and 35 percent, respectively. If the Rare UB is set to 5 percent, both road sections A and B lie in the “Very-Severe & Occasional” cell of Table 2 and receive the same deduct value of 38. In this case, both road sections...
Table 7. Priority profiles corresponding to the PM, CM, and RM obtained from optimization

<table>
<thead>
<tr>
<th>Scenario</th>
<th>PM priority</th>
<th>PM std</th>
<th>CM priority</th>
<th>CM std</th>
<th>Rare UB threshold</th>
<th>Occasional UB threshold</th>
<th>LDR threshold</th>
<th>Optimal utility function</th>
<th>Base case utility function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base case scenario</td>
<td>7</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0.10</td>
<td>0.5</td>
<td>85</td>
<td>7.38e+8</td>
<td>7.39e+8</td>
</tr>
<tr>
<td>Low budget ($200,000/year)</td>
<td>7</td>
<td>1</td>
<td>1.92</td>
<td>1.16</td>
<td>0.107</td>
<td>0.5</td>
<td>86.31</td>
<td>7.00e+8</td>
<td>7.05e+8</td>
</tr>
<tr>
<td>Most likely budget ($340,000/year)</td>
<td>7</td>
<td>1</td>
<td>1.44</td>
<td>1.02</td>
<td>0.103</td>
<td>0.5</td>
<td>86.31</td>
<td>6.85e+8</td>
<td>6.91e+8</td>
</tr>
<tr>
<td>High budget ($400,000/year)</td>
<td>7</td>
<td>1</td>
<td>1.41</td>
<td>1.20</td>
<td>0.100</td>
<td>0.5</td>
<td>86.31</td>
<td>6.38e+8</td>
<td>6.46e+8</td>
</tr>
<tr>
<td>Extreme budget ($600,000/year)</td>
<td>7</td>
<td>1</td>
<td>0.57</td>
<td>1.00</td>
<td>0.093</td>
<td>0.5</td>
<td>86.31</td>
<td>6.38e+8</td>
<td>6.46e+8</td>
</tr>
</tbody>
</table>
have the same priority in terms of receiving limited funds for performing CM. This leads to a proportional allocation of the limited available budget among road sections A and B. On the other hand, if Rare UB is set to 10 percent, then the deduct values corresponding to road sections A and B will be equal to 16 and 38, respectively, meaning that road section B will have higher priority for receiving the limited maintenance funds. Furthermore, sensitivity analysis showed that the utility function is more sensitive to the Rare UB in comparison with the Occasional UB. As discussed earlier in this section, most road sections in all scenarios are repaired before they fall into a condition where they have major problems. Therefore, the Occasional UB remains relatively unimportant. Interestingly, the optimal Occasional UB threshold (see Table 7) found in this analysis closely corresponds to the threshold used by VDOT in practice, suggesting the wisdom of current practice.

Another insight is that LDR levels may be too coarse a metric to prioritize among different road sections in need of the same maintenance type (e.g. PM or CM). In essence, road sections with very different distress density levels (e.g. 11 percent and 45 percent) lie in the same cell of Table 2 and thus have the same LDR value. In order to give a higher priority to the road sections that are in a worse condition, the exact distress density levels of the road sections could be a more useful metric for budget allocation than LDR. The alternative metric will ensure that road sections with a higher distress density receive a higher priority for their required maintenance given the limited budget.

The optimization analysis shows little benefit in doing additional PM where currently “Do Nothing” is advised. Two mechanisms explain this observation. First, this policy increases maintenance demand, but does not add much value because it requires maintenance for road sections that are otherwise acceptable (i.e. road sections are in the “Not-Severe & Rare” and “Severe & Rare” conditions that have an LDR greater than the LDR threshold). This leads to a potential waste of resources even if enough resources are available to complete this maintenance operation. Second, the condition “Not-Severe & Occasional” where the impact of this policy change is mostly observed does not happen very often. The transition rate from the NS cracks to S cracks obtained through calibration (see Table 4) in most instances prevents the NS cracks from accumulating to the point where they become “Not-Severe & Occasional”. Instead, in most cases, the NS cracks become S cracks.

The analysis of the utility function with respect to the LDR threshold also shows little sensitivity to this parameter. In essence, it was found that this threshold is somewhat redundant. The maintenance operation for the cells that lead to LDRs greater than 85 (base case threshold is 85) is Do-Nothing (see Tables 2 and 5). Alternative priorities for road sections could be assigned with more precision when using the priority profiles and the Rare UB threshold. The LDR threshold therefore remains of limited use and its optimum value is relatively close to the current decision rules.

Finally, in comparing the base case with the optimized utility function results, it can be observed that current VDOT practices, while not optimal, are not terribly inferior to the suggested practices that were found through optimization. The major departure, i.e. the importance of PM vs. CM, only modestly impacts the utility function values. Overall, the base case rules of thumb that have evolved over the years in practice are fairly effective. This may not come as a surprise once we consider the length of time over which these rules have evolved and have been fine-tuned to the local conditions.
Conclusions

The contributions of this paper to the body of knowledge in system dynamics are twofold. First, this paper develops one of the first SD simulation models of the physics of road deterioration and maintenance. As such, it introduces new concepts and knowledge from the road deterioration into the SD literature. Second, this paper couples the conceptualization steps of the SD method with calibration and optimization to construct a concrete framework for designing effective maintenance policies. This framework increases the applicability of the SD approach for operational and tactical decision support in the field of maintenance. In fact, this combination of modeling, calibration, and policy optimization leverages the unique strengths of SD in conceptualizing and modeling managerially relevant problems while at the same time effectively using available numerical data. This combination has been useful and attractive for clients of this research project who are not knowledgeable in SD. We therefore hope that our analysis provides a blueprint for the successful introduction of SD methodology into many alternative problem domains.

The study also makes practical contributions relevant to highway maintenance. The policies that were found through optimization pointed to alternative priorities where preventive maintenance is preferred over corrective maintenance. The extent of this contribution is limited, however, by the small improvements that are feasible. In essence, the results suggest that if one follows the current decision-making strategy there is only limited room for improvement.

On the other hand, there might be alternative decision-making structures that can provide improvement opportunities. For example, the current “road condition evaluation procedure” described in the section on “Modeling the Dynamics of Road Deterioration” translates a density function distribution of road section severity levels into a single point in a $3 \times 3$ matrix (see Table 1). This simplification is valuable given its cognitive simplicity and drives much of the decision making and analysis conducted in practice. On the other hand, with increasingly computerized analysis and decision-making tools, there is an opportunity for using all the information available about the density and severity of distresses on the pavement to drive the optimum maintenance policy. Future research can further explore this area of work.

Additionally, this model can be augmented by incorporating a user-benefit module that could include vehicle operating costs, travel time costs, and safety to change the specification of the optimized utility function. This would provide the opportunity to analyze the effects of different maintenance alternatives on societal costs and benefits. More work to establish a comprehensive utility function is left for future extensions of the model.

The results presented in this paper are also limited by the number of road sections analyzed and the number of years for which data were available. Calibrating the model using a more comprehensive dataset of road sections, road conditions, and operational environments could add to the validity of the model and the usefulness of the results. Moreover, to be able to compare the optimization results with the real dataset, the optimization horizon was limited to the number of years for which data were available (i.e. 5 years). Considering that the developed model can potentially be run over a longer horizon, using a dataset that contains longer periods of time can further reinforce the “long-term perspective” in road maintenance planning. Finally, capturing the distress
generation and propagation process for other types of distresses (e.g. rutting) is an important future extension. Despite these limitations, in the long term, this line of research could contribute to a more efficient use of societal resources, greater level of maintenance services, and a highway and roadway system that is safer and more reliable.

**Note**

1. The weights associated with the Very-Severe, Severe, and Not-Severe distresses are assumed to have values of 3, 2, and 1, respectively.

**Acknowledgements**

The work described in this paper was funded in part by NSF grant # CMMI-0726789. The authors would like to express their sincere appreciation to Allen Williams and Jeff Wright of Virginia Department of Transportation (VDOT) for letting us tap into their wealth of knowledge. Any opinions, conclusions, and/or findings are those of the authors and do not necessarily reflect the views of NSF and/or VDOT.

**Biographies**

Saeideh Fallah-Fini is a Ph.D. student in Industrial and Systems Engineering at Virginia Tech. She received her bachelor's degree from the Isfahan University of Technology and her master's from the Sharif University of Technology in Industrial Engineering. Saeideh's research interests include performance measurement and productivity analysis.

Hazhir Rahmandad is an assistant professor of Industrial and Systems Engineering at Virginia Tech. He received his Ph.D. from MIT System Dynamics group and his undergraduate degree in Industrial Engineering from the Sharif University of Technology. Hazhir's research interests include organizational learning, dynamics of product development capability, comparing different simulation methods and model estimation methodologies.

Konstantinos (Kostas) Triantis is a professor of Industrial and Systems Engineering, the 2008–2010 R. H. Bogle faculty fellow, and an adjunct professor of Civil and Environmental Engineering at Virginia Tech. He is the director of the Systems Engineering Program and the co-director of the System Performance Laboratory (SPL). He received his undergraduate and graduate engineering degrees in Industrial Engineering from Columbia University.

Jesús M. de la Garza is the Vecellio Professor of Construction Engineering and Management at Virginia Tech and leads the Center for Highway Asset Management ProgramS (CHAMPS). Dr de la Garza has been appointed to the National Research Council's Board on Infrastructure and the Constructed Environment and has served as Director of the Information Technology and Infrastructure Systems program at the National Science Foundation. Dr de la Garza received his M.S. and Ph.D. from the University of Illinois and his B.S. in Civil Engineering from the Tecnologico de Monterrey.
References


Chasey AD. 1995. A framework for determining the impact of deferred maintenance and/or obsolescence of a highway system. Ph.D. dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.


Chowdhury T. 2007. Supporting document for pavement models and decision matrices development process used in the needs-based budget. Virginia Department of Transportation Asset Management Division, Richmond, VA.


