Pavement Network Maintenance Optimization Considering Multidimensional Condition Data

Kenneth D. Kuhn[1]

Abstract: Pavement management systems inventory historical and current conditions of roadway networks, predict the future conditions of such networks, and suggest schedules for maintenance, repair, and rehabilitation activities. Such systems typically rely on a composite condition index, a one-dimensional and often discrete measure of the overall structural health and/or serviceability of pavement. The index is used during deterioration modeling, user and agency cost estimation, and selection and scheduling of maintenance activities. Pavement can suffer from a large number of related but distinct distresses. Difficulties associated with unobserved heterogeneity have hampered efforts to accurately model deterioration through composite condition indexes. At the same time, optimization techniques used to generate recommended maintenance plans have been shown both to be sensitive to deterioration model specification and to become computationally intractable as condition data increase. This research describes how a large network of related sections of pavement, each one of which may be plagued by a number of different distresses, can be managed without reducing condition data to a composite index. The use of approximate dynamic programming mitigates the curse of dimensionality that has haunted distinct Markov decision problem formulations of the maintenance optimization problem and limited their complexity. A computational study illustrates how the proposed approach leads to more sophisticated maintenance decision rules, which can be used to ensure the suggestions of pavement management systems more closely match engineering best practices. The use of multidimensional condition data can also yield more accurate deterioration models and cost estimates. The techniques introduced in this paper in the context of pavement management could easily be applied within any infrastructure management system. **DOI: 10.1061/(ASCE)IS.1943-555X.0000077.** *© 2012 American Society of Civil Engineers.*

CE Database subject headings: Pavement management; Infrastructure; Maintenance; Optimization; Data analysis.

Author keywords: Pavement management systems; Infrastructure management; Condition index; Maintenance optimization.

Introduction

Data related to numerous potential functional deficiencies are collected for different sections of pavement and stored in databases within pavement management systems. Examples of the described data include the percentage surface area of a roadway covered in cracks and the depth of ruts in vehicle wheel paths. The volume of data being collected is growing rapidly, in part due to the emergence of automated, often high-speed pavement distress survey techniques [see, for instance, Tighe et al. (2008)]. The German authorities have been regularly using high-speed monitoring systems to collect multidimensional data on their federal road network for at least 15 years (Burger et al. 1994). Data are often combined into a single composite index representing the overall condition of a section of pavement. The composite condition index is then used during deterioration modeling, user and agency cost estimation, and suggested maintenance action selection within pavement management systems.

A recent study found that composite indexes purporting to measure the same attributes of condition rated test sections of pavement substantially differently (Gharaibeh et al. 2010). This conclusion is troubling given the proliferation of different composite condition indexes in use today. Difficulties arise even when a single condition index is used in the context of deterioration modeling. When sections of pavement with similar condition ratings are subject to similar stresses, future condition ratings can vary considerably (Carnahan et al. 1987). Models that seek to capture stochasticity have been and are being developed. However, it has proved difficult to select appropriate forms and to accurately parameterize such models. The problem is partly caused by the limited volume of available data tracking condition index evolution over time, and partly by the fact that the same change in condition index may reflect many different physical changes to the pavement. This has led to a distinction between stochasticity and model uncertainty, with the latter continuing to haunt pavement management systems (Kuhn and Madanat 2005). One of the major causes of stochasticity and model uncertainty in modeling composite condition index evolution is unobserved heterogeneity (Prozzi and Madanat 2003). Sections of pavement that are identical in terms of data analyzed currently are, in fact, behaving quite differently. There will always be some unobserved heterogeneity in what might be termed facility type, so that actual underlying deterioration processes differ, but what is really being discussed is unobserved heterogeneity in facility condition.

This and related problems have been studied previously, most notably through approaches that treat the actual condition of a facility as an unobservable, latent variable related to observable data. Research on condition as a latent variable is surveyed subsequently in this introduction, but note that in these research efforts, data on distinct distresses like rutting and cracking are combined to estimate the (not directly observable and one- or possibly two-dimensional) true condition of a section of pavement that is then used for cost estimation and maintenance optimization.

[1]RAND Corporation, 1776 Main St., Santa Monica, CA; formerly, Dept. of Civil and Natural Resources Engineering, Univ. of Canterbury, Private Bag 4800, Christchurch, New Zealand. E-mail: kkuhn@rand.org

Note. This manuscript was submitted on February 2, 2011; approved on September 8, 2011; published online on September 10, 2011. Discussion period open until May 1, 2013; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Infrastructure Systems***, Vol. 18, No. 4, December 1, 2012. © ASCE, ISSN 1076-0342/2012/4-270-277/$25.00.

J. Infrastruct. Syst. 2012.18:270-277.

In order to minimize uncertainty and problems associated with unobserved heterogeneity, it would seem logical to separately consider all available data during deterioration and cost modeling, as well as maintenance optimization. In contrast to previous research, this research focuses on separately analyzing the proliferation and severity of different distresses within the context of pavement management systems, especially during network-level pavement maintenance, repair, and rehabilitation action optimization. The primary technical difficulty in avoiding the use of composite condition indexes throughout pavement management systems, and using multidimensional data in their place would have to be the computational complexity of optimization algorithms. Approximate dynamic programming is used to overcome challenges related to computational burden.

Composite Condition Indexes

Pavement can suffer from a large number of related but distinct distresses. Depressions may form in vehicle wheel paths, a process known as rutting. Fractures may develop on the pavement surface, which is known as cracking, in a variety of different ways. There may be variations on the surface of the roadway large enough to impact ride quality or raise vehicle operating costs, indicating roughness. There may be smaller variations associated with the macrotexture of the roadway surface, which can affect the skid resistance of the roadway in wet conditions. There may be even smaller variations associated with the microtexture of the roadway surface, which can impact skid resistance in dry conditions. Raveling, flushing, shoving, and spalling are additional potential distresses. When investigating the deterioration of a specific section of roadway and selecting an appropriate maintenance or rehabilitation activity, engineers will often consider separate pieces of data related to separate deficiencies. There have been and continue to be significant research efforts to build mechanistic and mechanisistc-empirical models describing how individual distresses evolve over time in a variety of contexts. For example, Henning (2008) models rutting and cracking on New Zealand state highways. Madanat et al. (1995) presents a methodology for modeling the initiation and progression of distresses.

At public agencies worldwide, data regarding different distresses are combined into composite condition indexes. Some commonly used examples include the pavement condition index (PCI), the present serviceability index (PSI), the pavement quality index (PQI), the pavement overall index (POI), and the distress manifestation index (DMI) (Golroo and Tighe 2009). These indexes allow concise representation of the overall condition of pavement. For instance, a particular section of road can be rated 4 out of 5 without having to mention, for example, the depth of ruts in vehicle wheel paths and the extent of cracking on the roadway surface. The composite indexes are perfect for providing a quick snapshot of the condition of a roadway network and to communicate with political decision makers and the general public.

Here, the distress manifestation index is taken as an example of a composite condition index. The distress manifestation index is used in Ontario, Canada, and is based on the results of visual condition surveys describing the progression of several distresses. Measurements of the severity and density of individual distresses are weighted to reflect their relative importance and summed. Table 1, based on data from a similar table in Tighe et al. (2008), shows the weighting factors associated with various distresses for various types of pavement.

Eq. (1), also from Tighe et al. (2008), shows exactly how the weighting factors are used to calculate the distress manifestation index:

$$\mathrm{DMI} = 10 \times \frac{G - \sum_{i=1}^{n} W_i(s_i + d_i)}{G} \quad (1)$$

where i = index of distress type; n = total number of distress types considered for the given pavement type; W_i = weighting factor of distress type i (Table 1); s_i = observed severity of distress type i; d_i = observed density of distress type i occurrence; and G = maximum value of weighted summation of distress severity and density observations for the given pavement type.

Most composite condition indexes, like the distress manifestation index, are based on a weighted summation of distress-specific evaluations. As was mentioned previously, different indexes will rate the relative health of different sections of pavement differently (Gharaibeh et al. 2010). According to the cited study, the disparity "can be attributed to differences in the distress types considered, weighting factors, and the mathematical forms of the indices" (Gharaibeh et al. 2010). Unfortunately, there are no clear correct mathematical forms, weighting factors, or even sets of distress types to survey when constructing a composite condition index.

Maintenance Policy Optimization

The condition indexes described are used a few different ways within pavement management systems. The uses are discussed in this paper in the context of maintenance policy optimization, a process that makes use of deterioration and cost models. An example of a representative approach to maintenance optimization is presented as follows.

Table 1. Surface Distress and Weighting Factors for Distress Manifestation Index Calculation

Asphalt concrete pavement	W	Portland cement concrete pavement	W	Composite pavement	W	Bituminous surface treatment pavement	W
Raveling	3	Raveling	0.5	Raveling	3	Raveling	3
Flushing	1.5	Polishing	1.5	Flushing	1.5	Flushing	2
Rippling or shoving	1	Scaling	1.5	Spalling	2	Streaking	1
Rutting	3	Potholing	1	Tenting or cupping	2.5	Potholing	1
Distortion	3	Joint cracking or spalling	2	Rutting	3	Rippling and shoving	2
Multiple cracking	1.5	Faulting	2.5	Joint failure	3	Rutting	3
Alligator cracking	3	Distortion	1	Distortion and settlement	1	Distortion	3
Meander midlane cracking	1	Joint failure	3	Longitudonal meander cracking	2	Longitudonal cracking	1
Transverse alligator	3	Longitudinal meander cracking	2	Transverse cracking multiple	1	Pavement edge breaking	2
Centerline alligator	2	Transverse cracking	2	Transverse joint reflective cracking	2	Alligator cracking	3
Pavement edge single/multiple	0.5	Sealant loss	0.5	Centerline cracking single	0.5		
Pavement edge alligator	1.5	Diagonal corner/edge cracking	2.5	Centerline cracking multiple	1.5		

J. Infrastruct. Syst. 2012.18:270-277.

In the example, a network of roadway pavements is to be managed over a finite-planning horizon. The condition of each section of pavement in the network is rated each year according to a discrete composite index. The data are aggregated and the overall condition of the network is itself a state in a discrete state space. One option would be for the network condition state to be defined as the condition state of each section of pavement in the network. Alternately, more aggregate or disaggregate data could be used to define the network condition state. For the purposes of this example, the key is that the network condition is characterized by selecting one state out of a discrete state space.

Each year, one action plan from a set of feasible maintenance action plans is performed on the network. The action plan describes maintenance actions to take on each separate section of roadway in the network. In most cases, budget and resource constraints define the set of feasible action plans and ensure that the decision problem is not separable into section-specific subproblems. There are also interdependencies in the deterioration rates of different sections of pavement due to common dependencies on weather, traffic, and construction conditions. Although these interdependencies have not been captured adequately in existing pavement management systems, they do point to the need for true network-level pavement management.

The social costs of the network being in any condition state with any maintenance action plan applied are taken as an input. There have been a handful of research efforts [notably Ben-Akiva and Gopinath (1995)] describing how to estimate social costs. These costs account for agency costs performing relevant maintenance and user costs the public incur using the facility. User costs can incorporate diverse considerations, for example, the delay to motorists caused by speed limit restrictions reflecting various roadway surface imperfections, accident risk (related to roadway surface macrotexture and microtexture) and vehicle ride quality (related to roughness). Future costs are discounted. The goal will be to minimize the sum of discounted costs over time, i.e., to find the social optimal maintenance policy.

Transition probabilities capture the probabilities that the network will be in different condition states 1 year after being in a given state with a given action applied. These values serve as the model of deterioration. Established mechanistic models of distress evolution are typically abandoned, and empirical models, often based on quite limited data sets, predominate.

Given these definitions, Bellman's equation is as in Eq. (2):

$$V_t(i) = \min_{a\in A}[c(i,a) + \alpha \sum_{j\in I} p(j|i,a)V_{t+1}(j)] \\ \forall\ t \in \{0,1,2,\dots T-1\}, \qquad i \in I \tag{2}$$

where i and j = indexes of (network) condition state; I = set of all (network) condition states; a = index of maintenance action plan; A = set of all feasible maintenance action plans; t = index of year number; T = length of the planning horizon in years; $c(i,a)$ = (single-year) social cost of being in condition state i and applying action plan a; $p(j|i,a)$ = probability of being in state j 1 year after being in state i with a applied; α = discount factor; and $V_t(i)$ = expected discounted future costs when in condition state i in year t.

Assume salvage values, terms representing the costs of ending the planning period with the network in a given condition state, are known for all condition states i in I. Set $V_T(i)$ terms equal to the salvage values. Then use Eq. (2) to find costs from year $T-1$ onward for all possible network condition states in I. Repeat the procedure for year $T-2$, and so on. For year 0, solve Eq. (2) given the current (assumed known) condition of the facility to provide optimal social costs through the planning horizon. The actions that have been chosen along the way to minimize costs provide the optimal maintenance policy, describing what maintenance action-to-take as a function of year and network condition.

The introduced formulation estimates social costs by summing $c(i,a)$ terms that use the condition index as their sole source of information regarding the infrastructure being managed. Similarly, deterioration is modeled as transitions between condition index states through $p(j|i,a)$ terms. Finally, the found optimal maintenance policy lists action-to-take as a function of condition index state (and year). Thus, it becomes important that condition indexes capture all the information relevant for three tasks: estimating agency and user costs, modeling deterioration indefinitely into the future, and selecting maintenance actions.

The computational burden associated with maintenance strategy optimization rises extraordinarily rapidly as the set I grows. For each combination of year, action plan, and network condition state, Eq. (2) requires summing over the set of network condition states. The set I likely grows extraordinarily rapidly as the description of each pavement section's condition becomes more detailed. A network condition state definition that describes each pavement section's condition requires m^n network condition states for a network of n sections, each of which may be described m ways. For example, there would be eight network condition states for a network of three pavement sections each of which may be categorized as being in good or bad condition. GGG, GGB, GBG, BGG, GBB, BGB, BBG, and BBB, where the conditions of the individual facilities are listed in some predetermined order and G (B) indicates a facility in good (bad) condition. If it was not necessary to uniquely identify the condition of each section of pavement for planning purposes but only to note how many sections were in each facility condition state, there would still be

$$\binom{m+n-1}{m}$$

possible network condition states. In 1982, the state of Arizona was reported to be responsible for 7,400 sections of pavement related by a common maintenance budget (Golabi et al. 1982). The approach to maintenance optimization outlined previously is an example of dynamic programming. Concerns regarding computational burden reflect the associated curses of dimensionality.

Past research efforts have yielded techniques for mitigating the curses of dimensionality. Golabi et al. (1982) use decision variables that describe the fraction of all facilities that are in different condition states. The size of the associated optimization problem is made independent of the number of facilities being managed (though still dependent on the number of condition states considered). Identifying clusters of facilities in similar condition states and ordering facilities in a clever way allows Childress and Durango-Cohen (2005) to describe the condition of a network of n facilities with m possible condition states in less than m^n network condition states. In a precursor to this study, Kuhn (2010) introduces the use of approximate dynamic programming in maintenance optimization.

Latent Performance Models

The prior research arguably most relevant to this paper concerns the construction and use of latent performance models. According to this line of research, engineers make various observations of individual distresses when evaluating the health of any infrastructure facility. Each observation is "probabilistically related to the true condition of the facility," which is not directly observable

(Madanat 1993). True condition is usually described in terms of a single latent condition variable, or two such variables representing structural and functional condition. Ben-Akiva and Ramaswamy (1993) construct models relating distress observations to the latent condition of pavement as well as models of latent condition deterioration. Madanat (1993) introduces a framework in which numerous observations of pavement condition are used to refine estimates of probabilities that the pavement is actually characterized as being in different discrete condition states. The framework is motivated by the early work of Humplick (1992) modeling pavement condition measurement error. Madanat and Ben-Akiva (1994) use the developed stochastic models of latent condition deterioration in the context of infrastructure maintenance and rehabilitation optimization. Ben-Akiva and Gopinath (1995) focus on estimating user costs associated with sections of pavement, in the latent condition framework.

Chu and Durango-Cohen (2007) construct state-space specifications of time-series models of infrastructure deterioration "consistent with the latent performance modeling approach." Durango-Cohen (2007) links such time-series models to maintenance and rehabilitation optimization. The approach can be thought of as a combination of two modules, one estimating the latent condition of a facility using a Kalman filter algorithm and another optimizing maintenance and rehabilitation. The use of continuous-valued condition and maintenance action variables allows the approach to use standard time-series and linear-quadratic programming techniques. In particular, the continuous-valued variables and overall framework mitigate the curses of dimensionality and allow for consideration of multidimensional condition data during infrastructure maintenance optimization (although examples to date have been based on one-dimensional latent condition data). At the same time, there is some discrepancy between the use of continuous-valued action variables and actual maintenance planning for individual sections of pavement in which decisions are often binary (do we chip seal or not? do we resurface or not?). There are a number of research papers by Durango-Cohen and coauthors based on time-series models and estimation of latent infrastructure condition, with Durango-Cohen and Tadepalli (2006) being notable for its discussion of "advanced inspection technologies" such as those that have recently come into widespread use for pavement management. The research framework within the works cited offers many nice properties, including a rigorous statistical framework for deterioration modeling, consideration of measurement errors in infrastructure facility condition surveys, and the ability to accept data sets containing omitted observations.

This work is differentiated from prior research regarding latent performance models in that there is no split between condition estimation and maintenance optimization in this work. Data on different distresses are not manipulated and combined to yield an estimate of a one- or two-dimensional latent condition state, but rather serve as the definition of condition itself.

Approximate Dynamic Programming Framework

The approach to maintenance optimization used in this paper is based on approximate dynamic programming (ADP) applied to pavement management, as introduced by Kuhn (2010). The main benefit of the approximate dynamic programming approach is that the approach allows the user to control the computational burden of the algorithm, both in terms of how much data need to be stored in memory and how many calculations are necessary during the optimization process.

Recall that using Eq. (2) to evaluate maintenance plan a in year t when the (network) condition state is i requires calculating $c(i,a) + \alpha \sum_{j\in I} p(j|i,a)V_{t+1}(j)$. In particular, $p(j|i,a)$ and $V_{t+1}(j)$ values for all i and j in I must be stored in memory while a summation across the set I must be taken. All this is necessary just for consideration of one state i and maintenance plan a in one year t. Again recall that the set I will be roughly of the size m^n, where m is the number of condition states for any one section of pavement and n is the number of sections of pavement being managed. In an ADP approach, the future costs of maintenance actions are approximated using functions specified a priori. It is typically only necessary to store the functions and weights attached to the functions in memory. Calculations typically involve multiplying the functions by their respective weights and summing the results. Because the functions are chosen by the user, the computational burden is controlled by the user. The ADP approach is summarized in this section of the current paper to clarify the points made previously and to provide context for the rest of the paper. [Much of this section is a restatement of material from Kuhn (2010), which contains a broader description and discussion of ADP applied to infrastructure management.]

Let $Y_t(i_1, i_2, \ldots, a)$ represent the value of observing the network in year t and having characterized the status of various distresses using the data sets $i_1, i_2, \ldots$ and having selected, but not performed, the maintenance action plan a. Y is similar to V in Eq. (2), and is known as the value function.

Estimating the Value Function

It is common in approximate dynamic programming to model the value function as a linear combination of basis functions (Powell 2007) as in Eq. (3):

$$\hat{Y}_t(i_1, i_2, \ldots, a) \equiv \sum_{s=1}^{S} w_s \phi_s(i_1, i_2, \ldots, a, t) \tag{3}$$

where $i_1, i_2, \ldots$ = set of values for (multidimensional) network condition data, s = index of basis function number; and S = number of basis functions considered. The function $\phi_s(i_1, i_2, \ldots a, t)$ is basis function number s, evaluated in year t when the network is characterized by condition data $i_1, i_2, \ldots$ and action a has been selected, w_s is the weight attached to basis function number s, and $\hat{Y}_t(i_1, i_2, \ldots, a)$ is a function that estimates the expected discounted future costs at year t given network condition data $i_1, i_2, \ldots$ and maintenance action plan a.

The basis functions are fixed and capture important attributes of the conditions of the pavement sections being managed, the actions selected, and the year of analysis. Further research is necessary to identify the most appropriate basis functions for network-level pavement management systems, but one example of a basis function that might prove useful would be a function that counts the number of lane-kilometer sections of pavement where there are ruts in vehicle wheel paths more than 15 mm deep.

Weights associated with basis functions may be fixed if there is sufficient expert knowledge, or may be learned, generally using simulation. Defining the estimate of the value function requirs only to set S parameters. This is considerably simpler than setting the value function for each combination of year and observed network condition data as required by Eq. (2). Simulation to set the w_s terms can make use of several established models describing how distinct condition data evolve at the same time. Unlike in distinct approaches to maintenance optimization, very few restrictions on the form of deterioration models are necessary. Considering complex models describing different distresses appearing and propagating according to distinct but interrelated mechanisms is not a problem. Data regarding the condition of a section of pavement need not be limited to a discrete, or even one- or two-dimensional, variable.

A particularly interesting research question involves determining how many and what type of simulation runs are required to estimate the value function. The answer will depend on both the confidence assigned to initial estimates of basis function weights and the variability of (simulated) future costs when a tuple of condition data, maintenance action plan, and year are given. It is difficult to say more with regard to the general case, but this topic will again be discussed in the context of computational studies presented subsequently in this paper.

Say that in year t of simulation run z, action plan a^t is taken when the pavement network is characterized by condition data $i_1^t, i_2^t, \ldots$. The model would then approximate the future costs of managing the pavement network to be the current estimate of $\hat{Y}_t(i_1^t, i_2^t, \ldots, a^t)$. The following (simulated) year finds the network in condition state $i_1^{t+1}, i_2^{t+1}, \ldots$ and action plan a^{t+1} is applied. It is preferable to refine w_s terms so that the estimate of the value function in the initial state moves closer to a new estimate based on what was observed during the simulation. To do so, an error term is defined through Eq. (4):

$$e_{t,z} \equiv c(i_1^t, i_2^t, \ldots, a^t) + \alpha \hat{Y}_{t+1}(i_1^{t+1}, i_2^{t+1}, \ldots, a^{t+1}) - \hat{Y}_t(i_1^t, i_2^t, \ldots, a^t) \tag{4}$$

where $e_{t,z}$ = difference between the estimate of future costs in year t of simulation run z and the best available approximation 1 simulated year later; and $c(i_1^t, i_2^t, \ldots, a^t)$ = (single-year) social cost of having the network in state $i_1^t, i_2^t, \ldots$ and applying action a^t.

It is possible to use a gradient-based approach to correct the estimate of the value function. Note that calculating the gradient is straightforward given the definition of the value function estimate, as in Eq. (5):

$$\nabla_w \hat{Y}_t[i_1^t, i_2^t, \ldots, a^t] = \begin{bmatrix} \phi_1(i_1^t, i_2^t, \ldots, a^t, t) \\ \phi_2(i_1^t, i_2^t, \ldots, a^t, t) \\ \vdots \\ \phi_S(i_1^t, i_2^t, \ldots, a^t, t) \end{bmatrix} \tag{5}$$

The distance to move when refining w_s terms is determined by a step size. Assume the step size varies by simulation run. Then each $\omega\sigma$ is updated according to Eq. (6):

$$w_s < -w_s + \gamma_z\, e_{t,z}\, \phi_s(i_1^t, i_2^t, \ldots, a^t, t) \quad \forall\ s \in \{1, 2, \ldots S\} \tag{6}$$

where z = index of simulation run number; and γ_z = step size during simulation run z.

Learning is being done online and it is only ever necessary to store in memory one set of S weights, two sets of condition data, and two actions (the condition data, and actions relating to the current and immediately previous year of the current simulation run). Memory requirements pose a significant challenge for dynamic programming approaches to infrastructure management and limit the scope of condition data considered. Approximate dynamic programming approaches to infrastructure management do not have the same difficulties associated with computer memory requirements. Here weights are being adjusted to reflect maintenance actions being taken. Thus, the estimate of the value function being developed reflects decision making during the simulation. The procedure outlined in this paper is a heuristic, one way to come up with a reasonable estimate of the value function Y. Further research to identify alternate approaches to value function estimation would be worthwhile.

Selecting Maintenance Actions

Decision making both during simulation runs, when the value function is being estimated, and in practice, when suggested maintenance action plans are being selected, proceeds the same way. The goal is to minimize social costs recognizing the limited resources available to public agencies. Here, this translates into selecting the action plan in the set of feasible plans that minimizes the value function. Mathematically, in a year t observing condition data $i_1, i_2, \ldots$, the optimization problem posed can be solved by Eq. (7):

$$\min_{a \in A} \hat{Y}_t(i_1, i_2, \ldots, a) = \sum_{s=1}^{S} w_s\, \phi_s(i_1, i_2, \ldots, a, t) \tag{7}$$

The set A is typically restricted by budget and resource constraints. For instance, it may only be feasible to reconstruct a certain length of roadway in one year. Kuhn (2010) notes that under such conditions, it is possible to choose basis functions that are both well suited to estimating the value function and ensure the optimization problem faced during maintenance action plan selection can be solved efficiently. More specifically, maintenance optimization problems are cast as multichoice multidimensional knapsack problems (MMKPs). MMKPs have been studied previously by experts in the field of operations research and efficient exact and approximate solution search strategies have been developed.

Computational Studies

Computational studies were run to demonstrate and explore a few of the points brought up previously. In the first study, simulations involving management of 1,000 sections of pavement were run. Ruts developed in the vehicle wheel paths and cracks formed and spread on the surfaces of the pavement. Only two potential distresses were considered for the sake of simplicity.

Cracking was modeled by using the work of Madanat et al. (1995), which is based on data collected by the World Bank over 7 years of study of highways in Brazil and is described in Paterson (1987). A joint discrete-continuous model was developed; cracking on a particular section of pavement at a particular time has either been initiated or not and, if initiated, progresses stochastically at a given rate. Variables found to be significant in modeling cracking include the thickness of pavement surfacing, the structural number of the pavement, and the traffic using the pavement. Given these terms, the probability of crack initiation was found to be given by Eq. (8):

$$\text{Prob}(\text{cracking initiation}) = \Phi[11.925 - 0.041h + 0.0000906h^2 - 2.198\hat{snc} + 1.871(traf/\hat{snc})] \tag{8}$$

where h = thickness of pavement surfacing; $\hat{snc}$ = estimate of the structural number of the pavement; $traf$ = traffic in number of million 80-kN-equivalent single-axle loads; and $\Phi[]$ = standard normal cumulative distribution function.

Given that cracking has been initiated, the expected percent of total area covered by cracks increases each year according to Eq. (9):

J. Infrastruct. Syst. 2012.18:270-277.

$$E[\text{cracking}\ (\%\ \text{area per year})|\text{cracking initiated}] = 6.41 - 0.787\hat{snc} + 1.645(\text{traf}/\hat{snc}) - 0.781\lambda \quad (9)$$

where λ = error correction term to be set to the standard normal probability density function divided by the standard normal cumulative distribution function when both are evaluated at the value used in Eq. (7).

Rutting was modeled using the work of Henning (2008), based on data collected from specially designated long-term pavement performance modeling sections of the New Zealand state highway system. In terms of the variables defined previously, the model predicts that pavements with a thickness of more than 150 mm will develop ruts that progress each year in terms of millimeters of added rut depth, according to Eq. (10):

$$\text{rut progression}\ (\text{millimeters per year}) = 9.94 - 1.38 * \hat{snc})(\text{traf}) \quad (10)$$

In the computational study described in this paper, all sections of pavement were assumed to be 200 mm thick and to have a structural number of 3. Yearly traffic volumes were described as random variables uniformly distributed between 10,000 and 20,000 80-kN-equivalent single-axle loads. There were only two options available to the responsible public agency regarding a particular section of pavement during a particular year. The pavement could be reconstructed, or left alone. All assumptions were made again primarily for the sake of simplicity. The goal was not to develop a highly realistic simulation of a pavement management system, but only to show some of the pitfalls of the common approach to infrastructure management.

There are multiple models of the costs users incur as a function of pavement serviceability or overall condition, but not as a function of the extent or severity of multiple distinct distresses like rutting and cracking. For further discussion on this topic, the reader is referred to a study by Ben-Akiva and Gopinath describing how data on cracking, rutting, and other distresses can be combined to form a functional performance measure that can then be used to estimate user costs (Ben-Akiva and Gopinath 1995). In the study run in this paper, it was assumed that ruts of a depth greater than 15 mm and cracks that covered more than 20% of the surface area of a section of pavement indicated substandard conditions. This assumption was based on discussions with traffic engineers in New Zealand. The likely results of having substandard pavement include decreased ride quality, decreased vehicle speeds, and increased delay. There may also be an increased accident risk associated with substandard pavement, depending on the severity of cracking and rutting as well as the strength of the statistical links between cracking, rutting, and other roadway surface imperfections. A cost 10 times as great as the cost of roadway reconstruction was incurred for every year a section of pavement was characterized as being substandard.

Two trials were run using the approximate dynamic programming framework for pavement management described previously. Both trials involved describing the condition of a section of pavement in a year as being in one of 10,000 possible states. In trial 1, cracking and rutting were each evaluated and assigned one of 100 possible ratings. Basis functions counted the number of sections of pavement in each combination of cracking rating range, rutting rating range, and action applied. For example, one basis function ϕ_s counted the number of sections of pavement with a rut depth of between 9.4 and 9.6 mm, between 9.25 and 9.50 surface area cracked, and that were being reconstructed. In trial 2, data regarding cracking and rutting were combined into a hypothetical composite index. Weighting factors of 7 for cracking and 3 for rutting were used (based loosely on Table 1), although it will be shown subsequently that the weights are largely irrelevant to the points the results bring up. Here basis functions counted the number of sections of pavement in condition index ranges either being reconstructed or left alone. Over the course of 100,000 simulation runs, value functions were estimated and optimal policies developed by comparing basis functions for sections being reconstructed versus being left alone. A total of 100,000 simulation runs were used to ensure there were several observations of sections of pavement in each of the 10,000 possible facility condition states. Basis functions can be interpreted as estimates of discounted future costs incurred when one additional section of pavement is in a condition state defined by cracking and rutting progression ratings with a particular action applied.

The entire study took less than 4 min when run in C on a standard personal computer. Fig. 1 shows a graphical representation of the optimal policies found.

The graph on the left-hand side of Fig. 1 represents the maintenance policy suggested after trial 1, and the graph on the right-hand side is that suggested after trial 2. In both graphs, the extent of rutting increases toward the right while the extent of cracking increases toward the bottom of the graph. Sections of the graphs colored black indicate pavement conditions that would trigger

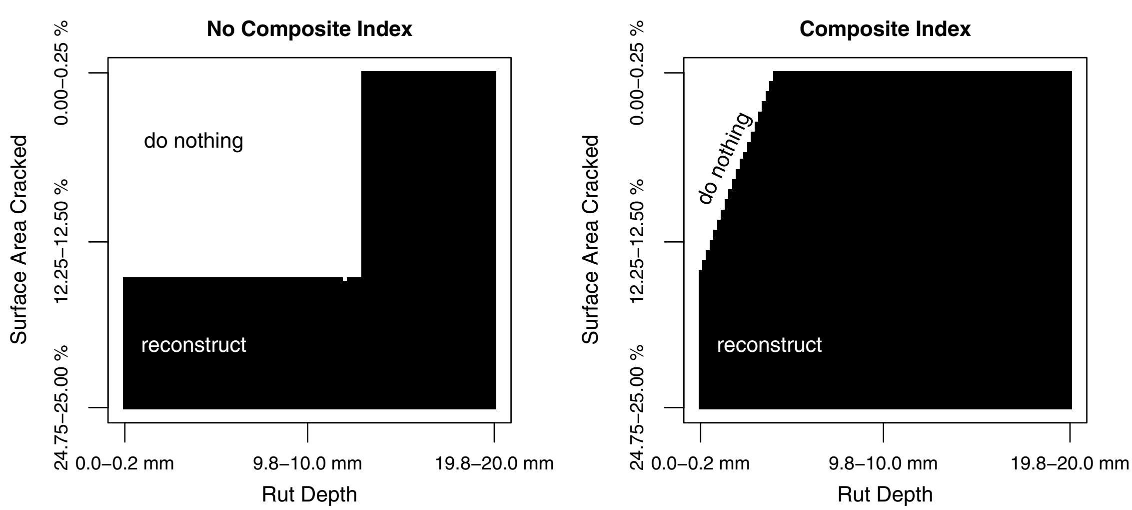

Fig. 1. Optimal policies with and without using a composite condition index

J. Infrastruct. Syst. 2012.18:270-277.

the recommendation of roadway reconstruction. Sections of the graph colored white indicate conditions that would not trigger such a recommendation. The small irregularity in the policy recommended when not using a composite condition index highlights the fact the value function was estimated using simulation and can include small irregularities. It also indirectly shows the level of detail with which it is possible to characterize the condition of a section of pavement when using an approximate dynamic approach to infrastructure management.

It is clear from Fig. 1 that in this computational study, the use of a composite condition index resulted in road reconstruction being recommended even when there was no danger of breaching the desired service standards. The condition index is unable to tell the difference between sections of pavement with worrying amounts of cracking but very little rutting and sections of pavement with less severe cracking but some rutting. Reconstruction is recommended for all sections of pavement where the condition index could have been caused by excessive cracking. There is a discrepancy between what engineers have defined as substandard and how the optimization framework has been set up. The one-dimensional composite condition index yields a one-dimensional decision rule. A straight line separates conditions in which reconstruction is recommended from conditions in which no maintenance is recommended. Altering the weights attached to rutting and/or cracking would alter the slope of the line, but not the form of the decision rule. The result is a recommended maintenance regime that does not match the prerogatives of engineers. There is the potential for public agencies to overspend on maintenance or for engineers to lose faith in the pavement management system. The actual optimal policy would be to reconstruct whenever either cracking or rutting reached its respective critical value at which there was a chance of reaching substandard conditions if the section was not reconstructed. The optimal policy is essentially the policy shown in the graph at left in Fig. 1, minus the irregularity. The value functions found indicated that the discounted future costs of managing a like new section of pavement were approximately 40% higher when using a composite condition index.

The study focused on the maintenance policy optimization function of pavement management systems. Within the study, agency costs of maintenance actions and user costs referencing pavement condition were assumed to be known. However, according to the study set up, accurate estimation of yearly social costs required knowledge of both the extent of cracking and the depth of ruts on a roadway surface. These are data a pavement management system employing a composite condition index would not typically utilize in yearly user and agency cost estimation.

In the first computational study, both approaches to the pavement management problem characterized pavement condition using a discrete state space. Theoretically, traditional dynamic programming approaches based on Eq. (2) could have been used. A second computational study was run to demonstrate why maintaining distinct distress-specific condition data sets in network-level pavement management systems necessitates the use of approximate dynamic programming. Again, there are sections of pavement to be managed that experience cracking and rutting according to the models previously introduced. A budget constraint made it impossible to reconstruct more than 10% of the sections of pavement in the network and ensured the problem was not separable. One experiment tracked the run time of a dynamic programming algorithm to find the optimal maintenance policy when 10 sections of pavement were to be managed for 30 years. Other experiments followed with networks of 11 and then 12 sections of pavement. The process continued until there were 112 sections of pavement to be managed and the required data

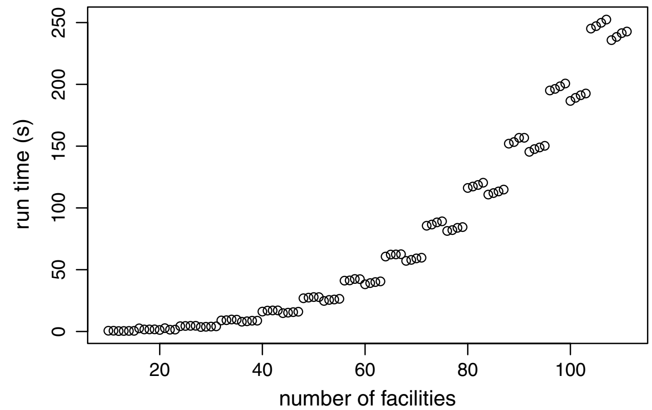

Fig. 2. Run time of dynamic programming algorithm as a function of the number of sections of pavement

could not be stored in the primary memory of the personal computer running the experiments. Fig. 2 shows the run times observed as a function of the number of sections of pavement being managed.

The results of the second computational experiment demonstrate that traditional dynamic programming approaches to infrastructure management do not scale well, making them unsuitable for use when managing large numbers of sections of pavement and keeping track of multidimensional, section-specific condition data. Note the exponential shape of the run times shown in Fig. 2, as well as the fact that a network of as few as 112 sections of pavement created unreasonable memory requirements.

Discussion

The relative importance of the points brought up in this paper depends on a few factors. The first of these factors is the extent to which the evolution patterns of distinct distresses are related. If cracking, rutting, flushing, shoveling, and spalling are all symptoms of one underlying deterioration process, then the use of a composite condition index makes some sense. Further research is needed to examine how interrelated distresses are from a statistical point of view. A (related) factor worth considering is the extent to which engineering decision making at the facility level, as well as roadway user costs, can be described using one-dimensional measures of pavement condition. Network-level pavement management systems for managers model decision making at the facility level done by engineers, and the fidelity of the model depends on how well the nuances of engineering judgment are captured. The techniques introduced in this paper could be used immediately to revamp decision making within network-level pavement management systems. The magnitude of the benefits of doing this are difficult to accurately estimate at this point.

The techniques introduced in this paper could actually be used in the context of any infrastructure management system. Management systems have been developed for bridges (Golabi and Shepard 1997), sewer systems (Wirahadikusumah et al. 2001), and countless other infrastructure networks. Many of these management systems refer to a single metric to describe the condition of an asset. The volume of data regarding infrastructure networks being collected is growing rapidly and management systems are recognized repositories for all available data. Thus, there is great potential for improving infrastructure deterioration and cost estimation models, as well as maintenance activity suggestions, within infrastructure management systems.

J. Infrastruct. Syst. 2012.18:270-277.

Conclusion

Much of the literature on optimizing the selection of pavement maintenance activities revolves around one-dimensional and discrete measures of overall pavement condition. Deterioration is modeled using probabilities describing how composite condition indexes evolve over time, but it has proven difficult to accurately establish such probabilities. In truth, pavement deteriorates as related but distinct distresses accumulate and propagate. Distress-specific, rather than overall condition, data are often collected. This work introduces a framework for separately considering the evolution of different distresses in a network-level pavement management system. Approximate dynamic programming is used to ease the computational burden during maintenance action policy selection, which would be too great if conventional approaches to infrastructure management were simply extended.

There appears to be a widely held belief that existing mechanistic and mechanistic-empirical models of how facilities deteriorate must be simplified during maintenance, repair, and rehabilitation strategy optimization. Composite condition indexes are likely useful in communications with politicians and the general public, but they are not required for maintenance optimization. All data available should be considered until such a point as some is found to be redundant when selecting maintenance action plans. This is what engineers do in practice when examining individual sections of pavement. Further work is needed to model costs associated with various distresses, to forecast the evolution of individual pavement distresses, and to establish basis function for use in approximate dynamic programming based infrastructure management.

References

Ben-Akiva, M., and Gopinath, D. (1995). "Modeling infrastructure performance and user costs." *J. Infrastruct. Syst.*, 1(1), 33–43.

Ben-Akiva, M., and Ramaswamy, R. (1993). "An approach for predicting latent infrastructure facility deterioration." *Transp. Sci.*, 27(2), 174–193.

Burger, W., Canisius, P. P., and Sulten, P. (1994). "Toward a new pavement management system in Germany: Organization, data collection, experiences, and innovations." *3rd Int. Conf. on Managing Pavements*, Vol. 1, Transportation Research Board (TRB) Committee on Pavement Management Systems, San Antonio, TX, 150–160.

Carnahan, J., Davis, W., Shahin, M., Keane, P., and Wu, M. (1987). "Optimal maintenance decisions for pavement management." *J. Transp. Eng.*, 113(5), 554–572.

Childress, S., and Durango-Cohen, P. (2005). "On parallel machine replacement problems with general replacement cost functions and stochastic deterioration." *Nav. Res. Logist.*, 52(5), 409–419.

Chu, C., and Durango-Cohen, P. (2007). "Estimation of infrastructure performance models using statespace specications of time series models." *Transp. Res. Part C*, 15(1), 17–32.

Durango-Cohen, P. (2007). "A time series analysis framework for transportation infrastructure management." *Transp. Res. Part B*, 41(5), 493–505.

Durango-Cohen, P., and Tadepalli, N. (2006). "Using advanced inspection technologies to support investments in maintenance and repair of transportation infrastructure facilities." *J. Transp. Eng.*, 132(1), 60–68.

Gharaibeh, N. G., Zou, Y., and Saliminejad, S. (2010). "Assessing the agreement among pavement condition indices." *J. Transp. Eng.*, 136(8), 765–772.

Golabi, K., Kulkarni, R., and Way, G. (1982). "A statewide pavement management system." *Interfaces*, 12(6), 5–21.

Golabi, K., and Shepard, R. (1997). "Pontis: A system for maintenance optimization and improvement of US bridge networks." *Interfaces*, 27(1), 71–88.

Golroo, A., and Tighe, S. (2009). "Fuzzy set approach to condition assessments of novel sustainable pavements in the Canadian climate." *Can. J. Civ. Eng.*, 36(5), 754–764.

Henning, T. (2008). "The development of pavement deterioration models on the state highway network of New Zealand." Ph.D. thesis, Univ. of Auckland, Aukland, New Zealand.

Humplick, F. (1992). "Highway pavement distress evaluation: Modeling measurement error." *Transp. Res. Part B*, 26(2), 135–154.

Kuhn, K. (2010). "Network-level infrastructure management using approximate dynamic programming." *J. Infrastruct. Syst.*, 16(2), 103–111.

Kuhn, K., and Madanat, S. (2005). "Robust maintenance and rehabilitation policies for a system of infrastructure facilities." *Transp. Res. Part C*, 13(5–6), 391–404.

Madanat, S. (1993). "Optimal infrastructure management decisions under uncertainty." *Transp. Res. Part C*, 1(1), 77–88.

Madanat, S., and Ben-Akiva, M. (1994). "Optimal inspection and repair policies for infrastructure facilities." *Transp. Sci.*, 28(1), 55–62.

Madanat, S., Bulusu, S., and Mahmoud, A. (1995). "Estimation of infrastructure distress initiation and progression models." *J. Infrastruct. Syst.*, 1(3), 146–150.

Paterson, W. D. O. (1987). *Road deterioration and maintenance effects: Models for planning and management*, World Bank, Washington, D.C.

Powell, W. (2007). *Approximate dynamic programming: Solving the curses of dimensionality*, Wiley, New York.

Prozzi, J., and Madanat, S. (2003). "Incremental nonlinear model for predicting pavement serviceability." *J. Transp. Eng.*, 129(6), 635–641.

Tighe, S., Ningyuan, L., and Kazmierowski, T. (2008). "Evaluation of semiautomated and automated pavement distress collection for network-level pavement management." *Transportation Research Record 2084*, Transportation Research Board, Washington, D.C.

Wirahadikusumah, R., Abraham, D., and Iseley, T. (2001). "Challenging issues in modeling deterioration of combined sewers." *J. Infrastruct. Syst.*, 7(2), 77–84.

J. Infrastruct. Syst. 2012.18:270-277.