Computation of Infrastructure Transition Probabilities Using Stochastic Duration Models

Rabi G. Mishalani[1] and Samer M. Madanat[2]

Abstract: Sound infrastructure deterioration models are essential for accurately predicting future conditions that, in turn, are key inputs to effective maintenance and rehabilitation decision making. The challenge central to developing accurate deterioration models is that condition is often measured on a discrete scale, such as inspectors' ratings. Furthermore, deterioration is a stochastic process that varies widely with several factors, many of which are generally not captured by available data. Consequently, probabilistic discrete-state models are often used to characterize deterioration. Such models are based on transition probabilities that capture the nature of the evolution of condition states from one discrete time point to the next. However, current methods for determining such probabilities suffer from several serious limitations. An alternative approach addressing these limitations is presented. A probabilistic model of the time spent in a state (referred to as duration) is developed, and the approach used for estimating its parameters is described. Furthermore, the method for determining the corresponding state transition probabilities from the estimated duration models is derived. The testing for the Markovian property is also discussed, and incorporating the effects of history dependence, if found present, directly in the developed duration model is described. Finally, the overall methodology is demonstrated using a data set of reinforced concrete bridge deck observations.

DOI: 10.1061/(ASCE)1076-0342(2002)8:4(139)

CE Database keywords: Computation; Infrastructure; Stochastic models; Deterioration.

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

Introduction

Infrastructure maintenance and rehabilitation decision making is based on current and future facility conditions. Current conditions are measured, and consequently their accuracy depends on the measurement technology. Future conditions, on the other hand, are predicted using a deterioration model. Therefore, accurate predictions are essential for effective maintenance and rehabilitation decision making. Infrastructure condition is often represented by discrete ratings or states. For example, for bridge decks, the Federal Highway Administration (FHWA) bridge rating system depicted in Table 1 is commonly used where inspectors employ ratings of 9 to 0, with 9 representing near-perfect condition (FHwA 1979). For pavement maintenance and rehabilitation decision making it is also common to use discrete condition states such as those in Golabi et al. (1982). The use of a discrete representation of facility condition makes it necessary to develop discrete-state models of deterioration.

The deterioration of an infrastructure facility is a stochastic process that varies widely with several factors, many of which are generally not captured by available data. Therefore, probabilistic models are often used to characterize deterioration. Two types of discrete-state probabilistic models have been used for infrastructure facility deterioration prediction: discrete-time, state-based models and time-based models. Mauch and Madanat (2001) described both types of models and discussed several examples from the infrastructure literature. Discrete-time, state-based models, such as Markov chains, characterize the probability that a facility undergoes a change in condition state at a given discrete time, given a set of explanatory variables such as design attributes, traffic loading, environmental factors, age, and maintenance history. Time-based models, on the other hand, characterize the probability density function of the time it takes an infrastructure facility to leave a particular condition state once entered (this time is referred to as state duration), given the same set of explanatory variables. As Mauch and Madanat (2001) observed, it is possible to use one model to determine the dependent variable of the other. More specifically, condition state transition probabilities can be determined from the probability density function of state duration and vice versa.

Mauch and Madanat (2001) discussed the appropriateness of using each of the two types of models based on empirical considerations relating to the specific nature of the condition data available for model development and estimation. For example, the estimation of time-based models would be practically feasible if frequent observations over long periods of time are available. In this paper, the focus is on methodological considerations relating to the nature of the assumptions made in estimating each model type. More specifically, an argument is made in favor of adopting time-based models in estimating state-duration probability density functions from which the transition probabilities of discrete-time, state-based models can be accurately determined. This approach will become more widely feasible with the advent of advanced inspection systems such as global bridge monitoring, where it is becoming cost-effective to use in situ sensors that collect detailed data almost continuously over time.

[1]Assistant Professor, Ohio State Univ. Dept. of Civil and Environmental Engineering and Geodetic Science, 470 Hitchcock Hall, 2070 Neil Ave., Columbus, OH 43210. E-mail: mishalani@osu.edu

[2]Professor, Univ. of California, Dept. of Civil and Environmental Engineering, 114 McLaughlin Hall, Berkeley, CA 94720. E-mail: madanat@ce.berkeley.edu

Note. Associate Editor: Kenneth Maser. Discussion open until May 1, 2003. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on August 29, 2001; approved on March 5, 2002. This paper is part of the ***Journal of Infrastructure Systems***, Vol. 8, No. 4, December 1, 2002. ©ASCE, ISSN 1076-0342/2002/4-139–148/$8.00+$.50 per page.

Table 1. Concrete Bridge Deck Condition Ratings (FHwA 1979)

Rating	Condition Indicators (% deck area) Spalls	Delaminations	Electrical potentials	Chloride content (lb/cu yd)
9	None	None	0	0
8	None	None	None>0.35	None>1.0
7	None	<2%	45%<0.35	None>2.0
6	<2% Spalls or sum of all deteriorated or contaminated deck concrete<20%			
5	<5% Spalls or sum of all deteriorated or contaminated deck concrete 20 to 40%			
4	>5% Spalls or sum of all deteriorated or contaminated deck concrete 40 to 60%			
3	>5% Spalls or sum of all deteriorated or contaminated deck concrete>60%			
2	Deck structural capacity grossly inadequate			
1	Deck repairable by replacement only			
0	Holes in deck—danger of other sections of deck falling			

1. The contributions of this paper are thus fourfold, as follows: Formalize the development of stochastic discrete-state infrastructure deterioration models. Specifically, a time-based probabilistic model is developed, and the approach used for estimating the model parameters is described.
2. Formulate the mathematical relationship between time-based models and discrete-time, state-based models. Specifically, the limitations of existing state transition probability estimation methods are pointed out, and the method for determining these probabilities from the developed time-based state-duration models is derived.
3. Describe how different mechanistic assumptions, such as age heterogeneity and state dependence (i.e., the absence of the Markovian property), can be tested for statistically and captured by the developed time-based model.
4. Demonstrate the overall methodology using a data set of concrete bridge deck observations.

The paper is organized as follows. The next section discusses the limitations of the state-of-the-art methods for estimating state transition probabilities. The third section describes the development of a stochastic state duration model and presents the approach used for the statistical estimation of the model parameters. Moreover, the method is derived for determining the transition probabilities of a discrete-time, state-based model from estimated state duration models. The section following that discusses how various assumptions regarding history dependence are tested, and if found valid, incorporated in the model. The fifth section uses a data set from the National Bridge Inventory to demonstrate the concepts and methods described in the previous sections. Finally, the paper closes with a conclusion and discussion of the feasibility of applying the developed methodology.

Transition Probability Estimation Literature

Several researchers have addressed the problem of estimating transition probabilities of discrete-time, state-based deterioration models. However, these current methods suffer from serious limitations. Madanat et al. (1995) summarized and critiqued a common method for estimating Markov transition probabilities. This approach, referred to as the expected value method, minimizes a measure of distance between the theoretical expected value of the state and the state as predicted by a linear regression model (Carnahan et al. 1987; Jiang et al. 1988). The theoretical expected value of the state is derived from the structure of the Markov chain, and the linear regression model is estimated using observations of state as the dependent variable and age as the explanatory variable. As Madanat et al. (1995) pointed out, this approach suffers from several serious limitations. First, it does not explicitly capture the effect of various explanatory variables. Furthermore, the possible nonhomogeneity (i.e., time dependence) of the deterioration process can only be captured indirectly through ad hoc time segmentation. Moreover, the presence of an underlying continuous deterioration level is not recognized. Finally, the use of linear regression is inappropriate due to the ordinal nature of condition states.

In response to the limitations of the expected value method, a class of structured econometric methods such as Poisson regression and ordered probit models have been used to estimate the parameters of deterioration relationships from which transition probabilities are computed (Madanat and Wan Ibrahim 1995; Madanat et al. 1995, 1997). However, these models have their limitations as well. The most critical limitation is the assumption that the observed condition states are independent and identically distributed (this also applies to the expected value method). This assumption does not hold, however; since conditional on being made between two consecutive state transitions, state observations are equal to one another and hence deterministically related. Another limitation is the inconsistency between the cumulative nature of the assumed underlying continuous deterioration level and the noncumulative nature of the explanatory variables used. This inconsistency arises from the effect of censoring where the starting conditions are unavailable due to the lack of monitoring facilities upon construction and the discrete time nature of inspection.

Finally, DeStefano and Grivas (1998) estimated a time-based model motivated by the need for state transition probabilities. However, the time-based modeling approach has a serious limitation. A restricted specification is adopted such that the complex structure of the deterioration process reflecting the significant effects of the multitude of explanatory variables (such as those discussed in the introduction) is not captured.

The consequences of the limitations associated with the various methods discussed above are biased estimates of the state transition probabilities, as shown in Madanat et al. (1995). This will lead directly to poor predictions of future infrastructure condition, thus compromising maintenance and rehabilitation decision making. The methodology presented in this paper addresses all the limitations discussed here and therefore will allow for a better prediction of deterioration.

Methodology

General Concept

As already discussed, infrastructure deterioration models are mathematical relationships between a dependent variable, namely deterioration or change in condition, and a set of causal variables, including design attributes, traffic loading, environmental factors, age, and maintenance history. The challenge central to developing accurate deterioration models for infrastructure facilities is that condition is often measured on a discrete scale, such as inspectors' ratings. These ratings or states are indicators of facility con-

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

J. Infrastruct. Syst. 2002.8:139-148.

dition, and therefore the change in state in a given period is an indicator of deterioration over that period. However, facility deterioration itself is a continuous process. For example, in reinforced concrete bridge decks, reinforcing steel bar corrosion is gradual and slow. Furthermore, in the initial stages the deterioration processes are not observed directly because they occur at either the microscopic scale or the subsurface level. However, indicators of performance, such as inspectors' ratings, are observed. It is therefore necessary, in developing stochastic discrete-state deterioration models, to correctly account for both the underlying unobserved continuous state deterioration process and the observed discrete-state deterioration indicators.

The basic assumption of the model developed in this paper is the existence of a probabilistic relationship between the indicators of deterioration and the underlying deterioration process. For clarity of presentation and without loss of generality, the model is developed for the simple bivariate case where the facility can be in one of only two possible states, namely 1 and 0, with the latter state reflecting a poorer condition. The extension to the multivariate case is introduced when the derivation of the transition probabilities is discussed. In the bivariate case, the observed sequence of states for each facility takes the form 1,1, . . . ,1,0,0, . . . ,0. Once a facility has undergone a change in state from 1 to 0, then no further changes are possible because condition cannot improve without a rehabilitation activity, which is not relevant for a model strictly describing deterioration. Consequently, only the first part of the sequence ending with the first time that a state of 0 is recorded contains meaningful information. Furthermore, and as discussed in the previous section, conditional on being made between two consecutive state transitions, state observations are deterministically related and therefore cannot be assumed independent from one another, as the current state-of-the-art transition probability estimation methods do.

The facility-specific variables are defined as follows:

1. D_t = unobserved cumulative continuous deterioration level at time t;
2. $\mathbf{X}_t$ = vector of observed explanatory variables at time t; and
3. Y_t = discrete state condition rating at time t (which takes a value of either 1 or 0).

The time variable t is set to zero at the time of the most recent rehabilitation, or construction in case no rehabilitation is carried out. The variables D_t and Y_t are related to one another through a threshold k reflecting the boundary defining the condition states. Even though D_t decreases over increasing time t as the facility deteriorates, as long as the value of D_t remains above the threshold k, Y_t maintains a value of 1. Once the value of D_t drops below k, the value of Y_t changes to 0. In other words, a state of 1 is recorded at time t if and only if $D_t > k$, and, conversely, a state of 0 is recorded at time t if and only if $D_t \leqslant k$. The time at which $D_t = k$ represents the duration of state 1 and is denoted by $t = T$. Since deterioration is a stochastic process, D_t and Y_t are random variables, and therefore T is a random variable as well.

This paper develops a model of the duration T, this model is estimated using a series of observations of Y_t while recognizing the already-discussed nature of dependence among these observations. Furthermore, this model takes into account the effects of causal variables and allows for determining the probability a facility will undergo a state change at any time t. Thus, all the previously noted limitations in the current transition probability estimation literature are addressed.

Typically, discrete-time, state-based deterioration models are characterized by transition probabilities for points in time separated by a constant period Δ, which is usually set at one or two years, the commonly employed inspection periods. The probability of the transition out of state 1 at time t, therefore, is the probability of observing the facility having already undergone a drop in state (i.e., change in state from 1 to 0) at time $t+\Delta$ (i.e., $Y_{t+\Delta}=0$) conditional on an observed facility state of 1 at time t (i.e., $Y_t = 1$). This conditional probability is denoted by $R(t,\Delta)$ and is given by the following:

$$R(t,\Delta) = \mathrm{Prob}(t < T < t+\Delta \mid T > t)$$
$$= \frac{\mathrm{Prob}(t < T < t+\Delta)}{\mathrm{Prob}(T > t)} = \frac{F(t+\Delta) - F(t)}{1 - F(t)}$$
$$= \frac{F(t+\Delta) - F(t)}{S(t)} \tag{1}$$

where $F(t)$ = cumulative distribution function of the duration random variable T; and $S(t) = 1 - F(t)$, which is known as the survival function in the stochastic duration modeling literature. Hence, if $F(t)$ is known, $R(t,\Delta)$ can be computed for any value of t and Δ. In other words, given a time-based model characterizing the probability density function of the duration T, the transition probabilities of the corresponding discrete-time, state-based model can be readily determined from $R(t,\Delta)$.

Hazard Rate Function

The probability of the transition out of state 1, $R(t,\Delta)$, clearly depends on the period Δ. The larger the value of Δ, the higher the probability and vice versa. Therefore, it is convenient to construct a measure that reflects the average rate at which the transition out of a state will occur by dividing this probability by Δ. If the facility is monitored almost continuously (i.e., successive observations are taken at very short intervals), then the conditional probability of Eq. (1) divided by Δ reduces to the following:

$$\lambda(t) = \lim_{\Delta \to 0} \frac{R(t,\Delta)}{\Delta} = \lim_{\Delta \to 0} \frac{F(t+\Delta) - F(t)}{\Delta S(t)} = \frac{f(t)}{S(t)} \tag{2}$$

where $f(t)$ = probability density function (PDF) of T. In other words, based on Eqs. (1) and (2), the probability of a transition out of state 1 at time t over an infinitesimal period $\Delta = dt$ is $\lambda(t)\,dt$. The function $\lambda(t)$ is known as the hazard rate function in the stochastic duration modeling literature and reflects the instantaneous rate (or risk) at which a facility will transition out of its current state to the lower state after time t.

The hazard rate function plays a central role in the development of stochastic duration models, and therefore some additional related relationships are worth presenting. The hazard rate function, $\lambda(t)$, and the survival function, $S(t)$, are related to one another through the following:

$$\lambda(t) = \frac{-d \ln S(t)}{dt} \tag{3}$$

It is also useful to present the integrated hazard function given by the following:

$$\Lambda(t) = \int_0^t \lambda(\tau)\, d\tau \tag{4}$$

This implies the following regarding the survival function:

$$S(t) = e^{-\Lambda(t)} \tag{5}$$

For more detail on the above functions, relationships, and sto-

J. Infrastruct. Syst. 2002.8:139-148.

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

chastic duration models in general, see Greene (1997). As for their application to infrastructure deterioration, see Prozzi and Madanat (2000).

Duration Model Specification

The functional form of the hazard rate function, $\lambda(t)$, allows for a convenient interpretation of the nature of the phenomenon being modeled. A constant hazard rate function implies a process where the conditional probability of the transition out of the state—given by Eq. (1)—does not vary over time. This implies that no matter how long the time spent in a state might be, the probability of the transition out of that state remains constant and hence independent of that time. This reflects a process lacking memory, and it can be shown that the associated distribution of the duration T follows the exponential PDF. Such a characteristic is referred to as duration independence. A hazard rate function that is monotonically decreasing over time implies that the probability of the transition out of the state decreases with an increasing time spent in that state. Conversely, a hazard rate function that is monotonicaly increasing over time implies that the probability of the transition out of the state increases with an increasing time spent in that state. These characteristics are referred to as negative and positive duration dependence, respectively; see Greene (1997) for more detail on all three characteristics.

As pointed out earlier in this section, once the PDF, $f(t)$, of the duration T is known, the transition probabilities of the corresponding discrete-time, state-based models can be determined. Therefore, all that remains to complete the development of the methodology is the specification and estimation of $f(t)$. Given the intuitive interpretation of the hazard rate function, $\lambda(t)$, and its straightforward relationship with the probability density and survival functions given by Eq. (2), it is both convenient and appealing to specify the hazard rate function directly. Once this function is estimated, the PDF of duration can be determined, and consequently the transition probabilities can be readily computed.

Although material aging and fatigue imply that infrastructure deterioration is in general expected to follow a monotonicaly increasing hazard rate function reflecting positive duration dependence, in estimating such a model it is desirable to adopt a single functional form that can cover the spectrum of negative duration dependence, duration independence, and positive duration dependence. This would allow for the least restrictive specification where the results of the parameter estimation reveal the true nature of the process rather than imposing one from the onset. One such functional form corresponds to the duration T following the Weibull PDF. The hazard rate function associated with the Weibull PDF is given by the following:

$$\lambda(t)=\lambda p(\lambda t)^{p-1}=p\lambda^{p}t^{p-1} \tag{6}$$

where λ,p = parameters to be estimated.

The expected value of the duration T that follows the Weibull PDF is given by $E[T]=\lambda^{-p}$. Depending on the value that p takes, the process falls in one of the three duration dependence categories. If $0<p<1$, then $\lambda(t)$ is monotonicaly decreasing and reflects negative duration dependence. If $p=1$, then $\lambda(t)$ takes the constant value of λ and reflects duration independence, or lack of memory, where the Weibull PDF reduces to the exponential PDF. And if $p>1$, then $\lambda(t)$ is monotonicaly increasing and reflects positive duration dependence.

It is useful to point out that this explicit treatment of state duration in modeling deterioration reflects a special case of the semi-Markov process (Ross 1983) where the following holds:

1. Probability of entering the next state once a transition occurs in continuous time is one if that state is the next best condition state, and zero otherwise.
2. Time until this transition occurs follows the Weibull PDF.

Such a process does not possess the Markovian property due to the presence of duration dependence unless the Weibull PDF takes the form of the exponential PDF. In this special case, the semi-Markov process reduces to a continuous-time Markov chain (Ross 1983) where duration independence holds.

As already discussed, infrastructure deterioration depends on a set of causal or explanatory variables. Clearly, the Weibull hazard rate function of Eq. (6) does not reflect the effect of such exogenous variables. Although this is a serious limitation of this model, its extension to include these variables can be achieved simply through the replacement of the constant parameter λ with a function dependent on the relevant variables. To ensure that this function is strictly positive, the following exponential specification is adopted:

$$\lambda=e^{-\beta\mathbf{X}} \tag{7}$$

where $\mathbf{X}$ = column vector of exogenous variables (which includes the value one to capture the constant term); and β = row vector of parameters to be estimated. Substituting Eq. (7) in Eq. (6) results in the following exogenous variable-dependent hazard rate function:

$$\lambda(t)=e^{-\beta\mathbf{X}}p(e^{-\beta\mathbf{X}}t)^{p-1}=pe^{-p\beta\mathbf{X}}t^{p-1} \tag{8}$$

Under this more general specification, therefore, the duration is now a function of the values of p and $\mathbf{X}$. Note, however, that the nature of duration dependence is still strictly dependent only on p.

Estimation

The parameters p and β of Eq. (8) are determined using maximum likelihood estimation. In doing so, the effect of censoring must be accounted for. Censoring reflects the situation where the observed duration T is only known to be greater than or less than a certain limit (Allison 1998). In the bivariate case, the duration in state 1 can be measured accurately if the time of the facility's construction is known, and if inspections are made almost continuously over time and include the instance when the transition occurs. If, however, the construction time is unknown, then only a lower bound on duration will be known. This reflects a case of right censoring where T is only known to be greater than the time between the commencement of inspection and the transition. In the event of another scenario where the construction time is known but inspections are made less frequently, only an upper bound on duration will be known. This reflects a case of left censoring where T is only known to be less than the time between construction and the first instance the facility is observed to be in the poorer state.

The situation of infrequent (i.e., discontinuous) observations over time is common; however, not knowing the time of construction is not. Nevertheless, a right censoring situation does occur in the multivariate case when the transition into the currently observed state takes place prior to the commencement of inspection. Right censoring for this case will be discussed in more detail in the empirical analysis section. What is relevant for this discussion is that the effect of censoring can be readily accounted for in the formulation of the likelihood function to be maximized. Greene (1997) discusses the maximum likelihood estimation of duration

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

models using censored data in general, while Prozzi and Madanat (2000) and Mauch and Madanat (2001) discuss the same in the context of infrastructure deterioration applications.

Transition Probability Determination

As already discussed, once the duration model is estimated, the transition probabilities of the corresponding discrete-time, state-based model can be computed. In effect this amounts to determining the transition probabilities of a discrete-time, state-based process from an estimated semi-Markov process where time is continuous. In the bivariate case the probability of the transition out of state 1, which is given by $R(t,\Delta)$ of Eq. (1), is simply the probability of the transition from state 1 to state 0. Consequently, the transition probability from state 1 to state 1 is $1-R(t,\Delta)$. Under the specification of Eq. (8), where a Weibull PDF and the presence of exogenous variables are assumed, the integrated hazard rate function of Eq. (4) is given by the following:

$$\Lambda(t)=(\lambda t)^p \tag{9}$$

Hence, based on Eq. (5), the survival function takes the following form:

$$S(t)=\exp[-(\lambda t)^p] \tag{10}$$

Therefore, the probability of remaining in the same state, denoted by $P_{1,1}$, is given by the following:

$$P_{1,1}=1-R(t,\Delta)=\frac{\exp[-\lambda^p(t+\Delta)^p]}{\exp[-(\lambda t)^p]} \tag{11}$$

And the probability of the transition from state 1 to state 0, denoted by $P_{1,0}$, is given by the following:

$$P_{1,0}=R(t,\Delta)=1-\frac{\exp[-\lambda^p(t+\Delta)^p]}{\exp[-(\lambda t)^p]} \tag{12}$$

In Eqs. (9) through (12), λ is as given by Eq. (7). Thus, given an estimated Weibull hazard rate function of Eq. (8), the transition probabilities for the bivariate case can be computed based on Eqs. (11) and (12) for any point in time t and any period Δ.

For the multivariate case, the three-state scenario is considered. A duration model for each state is developed exactly as discussed under the bivariate case. However, given the presence of multiple duration models in the multivariate case, a subscript is introduced in the notation reflecting the state each of the duration models corresponds to. In this discussion, the three states are denoted by 2, 1, and 0 whereby the additional state 2 is introduced reflecting the situation where state 2 is the best condition state and 0 the poorest. Naturally, the probability of the transition from state 2 to state 2 over a period Δ, denoted by $P_{2,2}$, is as given by Eq. (11) but as it applies to state 2, and consequently is given by the following:

$$P_{2,2}=1-R_2(t,\Delta)=\frac{\exp[-\lambda_2^{p_2}(t+\Delta)^{p_2}]}{\exp[-(\lambda_2 t)^p]} \tag{13}$$

where $R_i(t,\Delta)$ = probability of the transition out of state i; λ_i = given by Eq. (7) for state i; and p_i = parameter p of Eq. (8) for state i.

The transition out of state 2 to state 1 within a period Δ, denoted by $P_{2,1}$, is the probability of observing state 1 at time $t+\Delta$ given that state 2 was observed at time t. For this to take place, the duration in state 1 has to be greater than the interval between the time the transition actually occurs and the end of the period Δ. If the transition occurs at some time τ between t and $t+\Delta$, the transition probability $P_{2,1}$ would be the product of two probabilities: the probability of the transition out of state 2 taking place at a time falling between τ and $\tau+d\tau$, and the probability that the duration of state 1 is greater than $t+\Delta-\tau$. Consequently, for any value of τ between t and $t+\Delta$, the transition probability is the integral of this product over all possible values of τ. Hence, the transition probability from state 2 to state 1 is given by the following:

$$P_{2,1}=\text{Prob}(\text{state}=1 \ \text{ at time } \ \tau=t+\Delta|T_2>t)$$
$$=\int_t^{t+\Delta}\text{Prob}(\tau|T_2<\tau+d\tau|T_2>t)\cdot\text{Prob}(T_1>t+\Delta-\tau) \tag{14}$$

where T_i = duration of state i. Based on Eq. (2), the first term of the integral of Eq. (14) is given by the following:

$$\text{Prob}(\tau<T_2<\tau+d\tau|T_2>t)=\frac{F_2(\tau+d\tau)-F_2(\tau)}{S_2(t)}=\frac{f_2(\tau)d\tau}{S_2(t)}$$
$$=\frac{f_2(\tau)d\tau}{S_2(t)}=\frac{\lambda_2(\tau)\cdot S_2(\tau)}{S_2(t)}d\tau \tag{15}$$

where $F_i(\tau)$ = cumulative distribution function of duration of state i; $S_i(\tau)$ = survival function of state i; $f_i(\tau)$ = PDF of duration of state i; and $\lambda_i(\tau)$ = hazard rate function of state i.

Based on the definition of the survival function, the second term of the integral of Eq. (14) is given by the following:

$$\text{Prob}(T_1>t+\Delta-\tau)=S_1(t+\Delta-\tau) \tag{16}$$

Substituting Eqs. (15) and (16) in Eq. (14) results in the following:

$$P_{2,1}=\int_t^{t+\Delta}\frac{\lambda_2(\tau)\cdot S_2(\tau)\cdot S_1(t+\Delta-\tau)}{S_2(t)}d\tau$$
$$=p_2\lambda_2^{p_2}\exp[(\lambda_2 t)^{p_2}]$$
$$\times\int_t^{t+\Delta}\tau^{p_2-1}\exp[-(\lambda_2 t)^{p_2}-\lambda_1^{p_1}(t+\Delta-\tau)^{p_1}]d\tau \tag{17}$$

As in Eq. (13), λ_i is as given by Eq. (7) for state i, and p_i is the parameter p of Eq. (8) for state i.

Finally, since all the transition probabilities sum to one, the transition probability from state 2 to state 0, denoted by $P_{2,0}$, is given by the following:

$$P_{2,0}=1-P_{2,2}-P_{2,1}=R_2(t,\Delta)-P_{2,1} \tag{18}$$

As in the bivariate case, the transition probabilities for the three-state multivariate case can be computed based on Eqs. (13), (17), and (18) for any point in time t and any period Δ. The extension of the above derivation to the case of more than three states follows a similar treatment. Nevertheless, when the discrete state model is based on a reasonably short period Δ, it is unlikely that more than two-step transitions will occur due to the deterioration over such a period. Therefore, in all likelihood the three-state treatment should suffice in practical applications.

Tests of Markovian Property

One of the most debated questions in the field of infrastructure deterioration modeling has been the validity of the Markovian

J. Infrastruct. Syst. 2002.8:139-148.

assumption. In the probability literature, this is equivalent to the assumption of independence from history, which has been referred to earlier as duration independence or lack of memory. In the bivariate case, history is characterized by the time already spent in the current state of 1. For the multivariate case, history is also characterized by a second dimension, namely age at the time the current state is entered. A number of researchers have attempted to directly incorporate dependence on history into discrete-time, state-based deterioration models (Madanat and Wan Ibrahim 1995; Madanat et al. 1995, 1997). In fact, dependence on the two dimensions of history can be formally tested for in the context of the developed methodology and, if found to hold, incorporated into the developed stochastic duration model in a natural manner.

First, the dependence on the time already spent in the current state is examined. If the transition probability out of the state is independent of the time already spent in the state, then the duration independence property applies and, as discussed in the previous section, the hazard rate function is constant reflecting a duration T that follows the exponential PDF. The dependence of the hazard function on the time already spent in the state can be identified by inspecting the shape of the estimated hazard rate function. In the case of the Weibull hazard rate function, this is simply reflected in the value of the parameter p. As already discussed, duration independence occurs when $p=1$, and positive duration dependence occurs when $p>1$. Hence, positive duration dependence can be tested for by testing the null hypothesis that the estimated value of the parameter p is equal to one. If this null hypothesis is rejected and the estimated value of p is greater than one, then positive duration dependence would be clearly present. If, on the other hand, the null hypothesis is not rejected, meaning that p is not statistically different from one, then the hazard rate function can be reestimated with p set to one reflecting the exponential PDF.

In examining the dependence on age for the multivariate case, it is important to clarify that age is defined as the time from the most recent rehabilitation, or construction if no rehabilitation has taken place, to the time the facility enters the current state. Hence, age is clearly different from the time already spent in the current state. Therefore, it is critical to include age as an explanatory variable in the specification of the hazard rate function of Eq. (8) if it is found to be statistically significant. Indeed, age dependence (also known as age heterogeneity) can be tested for by testing the null hypothesis that the estimated coefficient of the variable age in the hazard rate function is equal to zero. Again, if this null hypothesis is rejected, then age dependence is clearly present and the variable age should remain included as an explanatory variable in the duration model. If, on the other hand, the null hypothesis is not rejected, meaning that the coefficient is not statistically different from zero, then the variable age should not be included in the duration model (especially in the case where a large enough data set is being used).

Empirical Analysis

Data Description

The National Bridge Inventory (NBI) includes data on more than 475,000 bridges (excluding culverts) constructed of various materials, including concrete, wood, and steel (Washer and Nelson 1997). The Indiana Bridge Inventory (IBI) is a subset of the NBI. The data set used in this analysis reflects observations on a subset of the bridges covered by the IBI. Specifically, the data set consists of the reinforced concrete bridge deck observations. Table 2 lists the variables included in the data set and used in the estimation exercise discussed in this section. Bridge deck condition is assessed in accordance with the FHwA rating scheme. As already mentioned in the introduction section, the ratings range from states 9 to 0, where 9 reflects the best possible condition and 0 the worst. The definitions of these ratings are shown in Table 1. As for the temporal nature of the observations, the inspection period reflected in the data set is two years, with the earliest inspection made in 1974 and the most recent in 1984.

Table 2. Bridge Deck Variables Used in Analysis

Variable	Description
TIS	Time-in-state (time deck has been in current deck condition) (years)
State	Deck condition state associated with time-in-state (TIS)
Rcensor	Presence of right censoring:
	0 = False (i.e., uncensored)
	1 = True
Drop	Number of condition ratings deck dropped since last bridge inspection
Age	Age at beginning of time in state or at beginning of observation period (years)
ADT	Average daily traffic (vehicles)
ADTYr	Year of the ADT count
AvgADT	Average ADT (unweighted average of inspection record ADTs for bridge deck)
Region	Climatic region:
	0 = South
	1 = North
Type	Deck structure type:
	1 = Concrete
	2 = Concrete continuous
	5 = Prestressed concrete
	6 = Prestressed concrete continuous
HWClass	Highway system classification:
	1 = Interstate, rural, open to traffic
	2 = Interstate, urban, open to traffic
	3 = Other FA primary, rural
	4 = Other FA primary, urban
	5 = FA secondary rural, state jurisdiction
	7 = FA secondary rural, local jurisdiction
WearSurf	Wearing surface material and protective system:
	1 = Concrete, no protective system
	2 = Asphaltic concrete, no protective system
	6 = Asphaltic concrete with known membrane
	7 = Other
	9 = Asphaltic concrete with coated rebar protective system

As for maintenance and rehabilitation, ideally the estimated deterioration models and the corresponding transition probabilities should not reflect such activities, whose effects should be captured through separate transition probabilities. In this analysis, once a bridge deck is rehabilitated, a new set of ensuing state durations is considered. However, bridge decks may also have been subjected to routine maintenance, but such information is not available to identify such observations. As a result, the estimated models and the computed transition probabilities in this analysis exclude the effects of rehabilitation but do not exclude the effects of routine maintenance.

J. Infrastruct. Syst. 2002.8:139-148.

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

Table 3. Number of Censored and Uncensored Observations for Condition States 8 and 7

State	Total observations	Uncensored observations	Censored observations	Percent censored
8	368	107	261	71
7	1,092	219	873	80
Total	1,460	326	1,134	78

The dependent variable of interest is the time spent by a bridge deck in a given condition state. This variable is referred to as time-in-state, or TIS for short. The following approximations were adopted in extracting TIS from the observed condition rating over time:

1. When a condition state drops by one step between two consecutive inspections, the actual transition is approximated to occur exactly midway between the two corresponding inspection times.
2. When a condition state drops by 2 steps between two consecutive inspections, the TIS for the intermediate state is approximated to be one year.

As a result, TIS will either be right censored or uncensored. The second approximation above eliminates the occurrence of left-censored observations. Given that the inspections covered a span of 10 years, the majority of the TIS observations extracted from the data set are censored. Table 3 shows the number of uncensored and censored observations for condition states 8 and 7, the two states for which duration models are estimated and transition probabilities computed in this analysis. Age is defined as discussed in the previous section with the exception of a particular scenario. For the right-censored TIS observations where the time the state was entered is unknown (due to its occurrence prior to the commencement of inspection), the age at the beginning of inspection has to be used as an approximation instead of the age at the time the state was entered.

Finally, it is important to emphasize that the use of the FHwA rating scheme in this analysis is for demonstration purposes only. The developed methodology is not restricted to this scheme and applies to any discrete condition-state definition for any physical component of infrastructure facilities.

Estimation Results and Interpretation

The data set described above was used to estimate duration models for states 8 and 7 based on the Weibull PDF with exogenous variables—that is, as characterized by Eq. (8). Two separate models were specified and estimated for each state since the mechanisms of deterioration and the effects of the various causal factors vary over time, depending on the state the deck is in. For example, deterioration out of state 8 into state 7 is primarily attributed to chemical processes, which are reflected in the amount of chloride content. Deterioration out of state 7, on the other hand, is primarily attributed to mechanical processes, which are reflected in the amount of spalls on the concrete deck. As a result, the significant causal variables are not expected to be identical for each model. Furthermore, even in the case where a variable plays a role in both models, the magnitude of its influence as captured by its estimated coefficient may be different. Finally, the magnitude of the intercept may also be different.

Duration Models

Table 4 shows the parameter estimates for the duration model for state 8. All the estimated parameters are significant at the 1% level. Furthermore, the signs are intuitively correct. The age variable contributes to the reduction of the hazard rate through a positive estimated coefficient. Thus, the higher the age at the time the deck enters state 8, the more likely that the duration in that state will be longer. This is intuitive since a longer age at the time of state entry implies a slower rate of deterioration. The joint effect of severe weather and deicing is significant as well. The region variable takes a value of 1 if the deck is located in the north and hence is subjected to more severe winter conditions along with the common use of deicing salts, which contributes to the corrosion of concrete deck reinforcing steel bars. This variable is, then, expected to contribute to increasing the hazard rate (i.e., increasing the likelihood of a shorter duration in state 8), which is precisely the case, as indicated by the negative sign of its estimated coefficient.

Table 4. Parameter Estimates for State 8

Variable	Estimated parameter	t-statistic
Constant	2.13	9.44
Age	0.15	6.31
Region	−0.84	−7.06
Type2	0.38	3.09
HWClass1	−0.64	−2.89
HWClass3	−0.55	−3.62
HWClass5	−0.57	−3.36
WearSurf1	−0.80	−4.73
$1/p$	0.52	12.01

The estimation results indicate that a concrete continuous type structure reduces the hazard rate. As for highway class, the variables HWClass1, HWClass3, and HWClass5 take values of 1 if the deck belongs to various categories of rural roadways. The negative signs of the corresponding estimated coefficients reflect a contribution to an increase in the hazard rate. One possible interpretation of this result is that rural roadways are subject to lower design or maintenance standards than urban roadways. Finally, the estimation results indicate, through the estimated coefficient of WearSurf1, that a concrete wearing surface with no protective system contributes to increasing the hazard rate.

In addition to the effects of the causal variables discussed above, the value of the parameter p is estimated at 1.93. Furthermore, the 95% confidence interval is between 1.62 and 2.25, thus indicating that p is significantly different from 1 at the 5% level. As discussed in the previous two sections, this results in a monotonicaly increasing hazard rate function reflecting positive duration dependence, thus confirming, along with the significance of the coefficient of the age variable discussed above, that the deterioration process once a deck is, in state 8 does not possess the Markovian or lack-of-memory property. This result is graphically depicted in Fig. 1, which shows a clearly increasing hazard rate function for state 8 (up to a time-in-state of at least 50 years) of a representative deck with quantitative variables equal to the mean of the population of decks in state 8 and categorical variables equal to the mode.

Table 5 shows the estimated parameters of the duration model for state 7. Again, all the estimates are significant at the 1% level. Among the differences of interest in comparison with those of state 8, the coefficient of the age variable for state 7 was not found to be statistically different from zero. This suggest that once the mechanical aspect of deterioration takes effect, the age at the time state 7 is entered has no effect on the hazard rate, and hence the duration. However, the effect of traffic loading as cap-

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

J. Infrastruct. Syst. 2002.8:139-148.

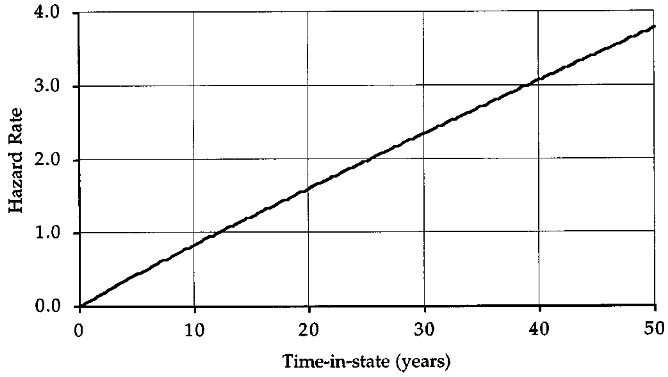

Fig. 1. Hazard rate function for state 8

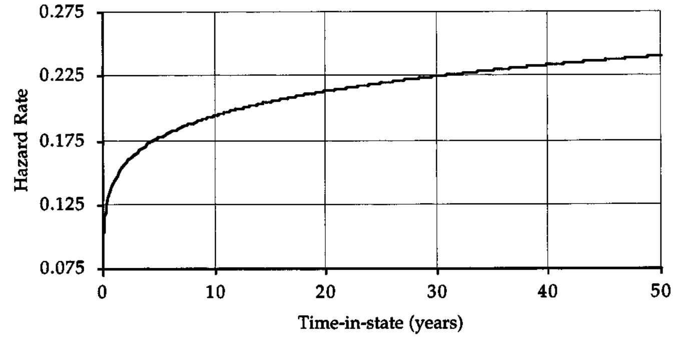

Fig. 2. Hazard rate function for state 7

tured by the AvgADT variable was found to be significant in contrast to the model for state 8. This influence is reflected in an increase in the hazard rate. This is consistent with the fact that a deck in state 7 deteriorates out of that state primarily as a result of mechanical rather than chemical deterioration processes where traffic loading is expected to play an important role. Furthermore, although the effect of highway class is similar to that found for state 8, the magnitude of the influence is stronger, as reflected in the larger absolute values of the estimated coefficients for variables HWClass1, HWClass3, and HWClass5. This could also be attributed to the primarily mechanical nature of deterioration once in state 7. Finally, the asphaltic concrete wearing surface with no protective system has an effect similar to that of the concrete wearing surface with no protective system, as reflected in the coefficient of the variable WearSurf2.

As for the question of duration dependence, although the estimated value of p at 1.13 is greater than one, it is not that much larger than one, as in the case of the model for state 8. In fact the 95% confidence interval is between 0.94 and 1.33, thus indicating that p is not significantly different from 1 at the 5% level. This is also reflected graphically in Fig. 2, where although the hazard rate function of the representative deck is still increasing, the range of variation is small in comparison to that of state 8, and the slope flattens out relatively quickly as the time-in-state increases. This, along with the lack of significance of the age variable, suggests that for state 7 the Markovian property might hold. If using a larger data set where inspections are conducted over a span of more than 10 years reveals similar results, then a revised duration model would be estimated where the Markovian property is imposed through the setting of p to one. As already discussed, this reduces the Weibull PDF to the exponential PDF. The difference between states 8 and 7 in terms of the nature of duration dependence emphasizes the value of allowing the estimation results to reveal this nature rather than imposing some assumption prior to the estimation.

Table 5. Parameter Estimates for State 7

Variable	Estimated parameter	t-statistic
Constant	5.07	11.30
AvgADT	-2.64×10^{-5}	-2.70
Region	-0.83	-5.57
HWClass1	-1.02	-3.87
HWClass3	-1.30	-5.15
HWClass5	-1.14	-4.57
WearSurf1	-0.99	-3.74
WearSurf2	-1.15	-4.23
$1/p$	0.88	11.26

Transition Probabilities

Once the deterioration process is modeled using a time-based formulation with all issues discussed in the literature review section addressed, as empirically demonstrated above, the transition probabilities for the corresponding discrete-time, state-based formulation can be accurately computed. In the methodology section, the transition probabilities for the bivariate case were derived and are given by Eqs. (11) and (12). This will suffice in computing the transition probabilities for state 7 where Eq. (11) for state 7 represents the transition from state 7 to state 7, while Eq. (12) represents the transition from state 7 to state 6 or lower. This is because models for state 6 and lower have not been estimated. However, in the case of state 8, it is possible to compute transition probabilities for state 8 to state 8, state 8 to state 7, and state 8 to state 6 or lower. In this case the transition probabilities for the three-state multivariate case derived in the methodology section applies where Eq. (13) for state 8 represents the transition from state 8 to state 8, Eq. (17) represents the transition from state 8 to state 7, and Eq. (18) represents the transition from state 8 to state 6 or lower.

To demonstrate the computation of the transition probabilities from the estimated duration models for states 8 and 7, these probabilities are computed for values of TIS up to 50 years for a period Δ of one year. Note that, although the data set used is based on a two-year inspection period, the developed methodology allows for computing the transition probabilities for any

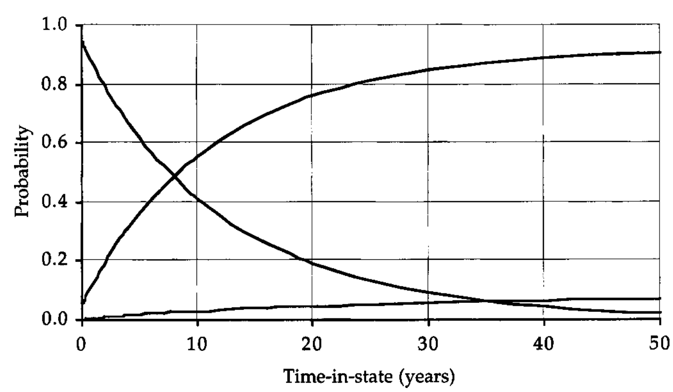

Fig. 3. Transition probabilities for state 8 ($P_{8,8}$=decreasing line, $P_{8,7}$=top increasing line, $P_{8,6\ \mathrm{or\ lower}}$=bottom increasing line)

J. Infrastruct. Syst. 2002.8:139-148.

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

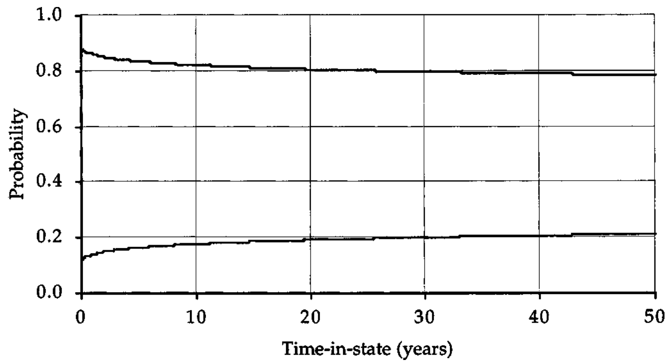

Fig. 4. Transition probabilities for state 7 ($P_{7,7}$=decreasing line, $P_{7,6 \text{ or lower}}$=increasing line)

value of Δ. Fig. 3 shows the transition probabilities for state 8 where the integral of Eq. (17) is evaluated numerically. Notice that, as expected, the probability of remaining in state 8 decreases as the time-in-state increases, while the probabilities of the transitions from state 8 to state 7 and from state 8 to state 6 or lower increase. Moreover, these variations with an increasing time-in-state are significant. This clearly reflects the positive duration dependence discussed above for state 8. Fig. 4 shows the transition probabilities for state 7. Although the probability of remaining in state 7 is decreasing and the probability of the transition from state 7 to 6 or lower is increasing with increasing TIS, the ranges of variation are much smaller in comparison to those of state 8, and the slopes flatten out relatively quickly as the TIS increases. This is clearly consistent with the above discussion regarding state 7 in relation to the Markovian property.

Conclusion

This paper presents the development of a time-based discrete-state stochastic duration model that takes into account the effects of causal variables. The estimation approach is based on using a series of state observations and recognizing that, conditional on being made between two consecutive state transitions, state observations are equal to one another and hence deterministically related. Furthermore, the methodology for determining the transition probabilities of the corresponding discrete-time, state-based model from the duration model is developed. Finally, testing for the Markovian property is discussed, and incorporating the effects of history dependence, if found present, directly in the developed duration model is described. Therefore, the developed methodology addresses all the previously noted limitations in the current state transition probability estimation literature. As a result, it allows for arriving at better predictions of deterioration and, hence, more effective maintenance and rehabilitation decisions.

The developed methodology was demonstrated using a data set of reinforced concrete bridge decks. Two separate duration models were estimated for states 8 and 7 based on the Weibull PDF with exogenous variables including age, traffic loading, environmental factors, structure type, highway class, and wearing surface type. The results indicate the significance of such causal variables. Furthermore, the testing for history dependence is demonstrated. In the case of the model for state 8, where deterioration is primarily governed by chemical processes, the age variable was found be statistically significant, and the hazard rate function was found to be monotonicaly increasing, thus reflecting positive duration dependence. These results formally confirmed that the deterioration process once a deck is in state 8 does not possess the Markovian or lack-of-memory property. On the other hand, in the case of state 7, where deterioration is primarily governed by mechanical processes, the age variable was not found to be statistically significant and the hazard rate function was found to be relatively flat, suggesting that the Markovian property might hold. In either case, the empirical analysis also demonstrated the application of the estimated time-based duration models in accurately computing the transition probabilities for the corresponding discrete-time, state-based models. These results further supported the conclusions reached for the examined bridge decks regarding the Markovian property.

In summary, a methodological argument is made in favor of adopting time-based models for accurately determining the transition probabilities of the corresponding discrete-time, state-based models. Even though the developed methodology is demonstrated empirically, as Mauch and Madanat (2001) pointed out, estimating stochastic duration models would be practically feasible if frequent observations over long spans of time are available. With the advent of advanced inspection systems such as global bridge monitoring, where it is becoming cost-effective to use in situ sensors to collect detailed data almost continuously over time, the accurate computation of transition probabilities from estimated duration models is expected to become more feasible on a wide-scale basis.

Acknowledgments

This research was partly funded through a grant from the National Science Foundation and a seed grant from the Office of Research at The Ohio State University. Discussions with Carlos Daganzo of the University of California, Berkeley, were instrumental in pointing to the need for addressing the limitations in the literature. The estimation results were produced through the assistance of Wenhui Zhao, a graduate research assistant at The Ohio State University.

References

Allison, P. D. (1998). *Survival analysis using the SAS system: A practical guide*, SAS Institute, Carey, N.C.

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

DeStefano, P. D., and Grivas, D. A. (1998). "Method for estimating transition probability in bridge deterioration models." *J. Infrastruct. Syst.*, 4(2), 56–62.

Federal Highway Administration (FHwA). (1979). *Recording and coding guide for structure inventory and appraisal of the nation's bridges*, U.S. Dept. of Transportation, Washington, D.C.

Golabi, K., Kulkarni, R. B., and Way, G. B. (1982). "A statewide pavement management system." *Interfaces*, 12(1), 5–21.

Greene, W. (1997). *Econometric analysis*, Macmillan, New York.

Jiang, Y., Saito, M., and Sinha, K. C. (1988). "Bridge performance prediction model using the Markov chain." *Transportation Research Record 1180*, Transportation Research Board, Washington, D.C., 25–32.

Madanat, S. M., Karlaftis, M. G., and McCarthy, P. S. (1997). "Probabilistic infrastructure deterioration models with panel data." *J. Infrastruct. Syst.*, 3(1), 4–9.

Madanat, S. M., Mishalani, R., and Wan Ibrahim, W. H. (1995). "Estimation of infrastructure transition probabilities from condition rating

J. Infrastruct. Syst. 2002.8:139-148.

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

data.'' *J. Infrastruct. Syst.*, 1(2), 120–125.
Madanat, S. M., and Wan Ibrahim, W. H. (1995). ''Poisson regression models of infrastructure transition probabilities.'' *J. Transp. Eng.*, 121(3), 267–272.
Mauch, M., and Madanat, S. M. (2001). ''Semiparametric hazard rate models of reinforced concrete bridge deck deterioration.'' *J. Infrastruct. Syst.*, 7(2), 49–57.
Prozzi, J., and Madanat, S. M. (2000). ''Analysis of experimental pavement failure data using stochastic duration models.'' *Transportation Research Record 1699*, Transportation Research Board, Washington, D.C., 87–94.
Ross, S. M. (1983). *Stochastic Processes*, Wiley, New York.
Washer, G. G., and Nelson, G. (1997). ''Development of a bridge steel database.'' *Public Roads*, 60(3), 27–29.

Downloaded from ascelibrary.org by Eth-Bibliothek on 12/09/13. Copyright ASCE. For personal use only; all rights reserved.

J. Infrastruct. Syst. 2002.8:139-148.