International Journal of Architectural, Civil and Construction Sciences
ISSN: 2415-1734
Vol:6, No:1, 2012

A Hidden Markov Model for Modeling Pavement Deterioration under Incomplete Monitoring Data

Nam Lethanh and Bryan T. Adey

Abstract**—**In this paper, the potential use of an exponential hidden Markov model to model a hidden pavement deterioration process, i.e. one that is not directly measurable, is investigated. It is assumed that the evolution of the physical condition, which is the hidden process, and the evolution of the values of pavement distress indicators, can be adequately described using discrete condition states and modeled as a Markov processes. It is also assumed that condition data can be collected by visual inspections over time and represented continuously using an exponential distribution. The advantage of using such a model in decision making process is illustrated through an empirical study using real world data.

Keywords**—**Deterioration modeling, Exponential distribution, Hidden Markov model, Pavement management

I. Introduction

A. Insufficient data

The prediction of future condition is necessary in the determination of optimal pavement intervention strategies (OISs). Accurate prediction requires that prediction models be validated using past condition data. In reality, however, the quantity and quality of past condition data is often insufficient to validate the prediction models and therefore to determine the OIS. Two examples of insufficient data from which to determine OISs are:

- There exists extensive data on the visual appearance of the pavement surface, e.g. percentage of surface area that is cracked, but there is sparse data on the longitudinal and transversal unevenness, friction or load bearing capacity of the road section.
- There exists extensive data on the road roughness, e.g. the International Roughness Index (IRI), but there is sparse data on the percentage of surface area that is cracked.

Determining OISs in either one of these cases may result in the exclusion of certain types of interventions, as appropriate triggers cannot be assigned that would allow interventions of certain types to be selected. In the first case, the optimal time to execute an intervention that alleviates a longitudinal unevenness problem cannot be determined by knowing only the evolution over time of the cracking of the pavement surface. In the second case, the optimal time to execute an intervention that improves pavement friction cannot be determined by knowing only the evolution over time of the roughness of the road [1].

The reasons for insufficient data include the lack of sufficient resources to conduct the necessary inspections, e.g. to purchase required equipment or to hire an adequate number of inspectors, and a perceived negative net benefit, i.e. it is not worth the effort to acquire the additional data.

B. Modeling deterioration and determining optimal intervention strategies:

The modeling of pavement deterioration in pavement management systems (PMS) in recent years has predominately been with the Markov model [2, 3, 4]. This trend is also evident in research on the management of other infrastructure objects, such as bridges [4, 5, 6]; and pipelines [7, 8]. Two advantages of Markovian models are:

- they allow generalization of the deterioration process into the transition pattern among condition states, which is suitable for representing pavement performance;
- they can be used in the absence of historical data, as the probability of observing future state depends only on the probability of observed condition states at the present. Thus, with a minimum of two visual inspections, deterioration progress of the pavement can be predicted.

With the Markov model, the deterioration of road sections is described using probabilities of transition between discrete condition states. The probabilities of being in each condition state in the Markov transition matrix multiplied by the predetermined or initial condition state probabilities give an overall assessment of pavement condition in respective time frames. In developed countries such as Japan, Switzerland and the United States of America, the definition of condition states (composite pavement indexes) have been standardized to facilitate pavement management. For example, the Pavement Condition Index (PCI), developed by the United States of Army Corps of Engineers [9] or the Management Criteria Index (MCI) used in Japan [10]. These composite pavement indexes are normally calculated using a deterministic relational function between several important pavement condition indicators, such as cracking, rut depth and roughness.

PMSs use these deterioration models to determine OISs, consisting of the condition state of the road section that is to trigger an intervention and the type of intervention to be determined. When inaccurate deterioration models exist, the future condition state of the road section, i.e. the value of the aggregate composite index, can not be accurately estimated, and therefore neither can the OISs.

A potential way to overcome this problem is to estimate the transition probabilities using a hidden Markov model (HMM).

Dr. Nam Lethanh is with the Institute of Construction and Infrastructure Management, Swiss Federal Institute of Technology (ETHZ), 8093 Switzerland, email: lethanh@ibi.baug.ethz.ch

Prof. Bryan T. Adey is with the Institute of Construction and Infrastructure Management, Swiss Federal Institute of Technology (ETHZ), 8093 Switzerland, email: adey@ibi.baug.ethz.ch

HMMs have been significantly studied and broadly applied in the field of image processing and speed recognition [11]. There has, however, been very little investigation of the use of HMMs in the field of infrastructure management, and the research that has been conducted has focused principally on the elimination of measurement errors (e.g. [10]).

To use HMMs to model road deterioration it is necessary to assume that road deterioration is a hidden process, i.e. it cannot be observed directly, that can be deduced through the evolution of an observable process, such as increasing roughness. The transition probabilities for the Markov model are then estimated using the mathematical relationship between the hidden process and the observable process. The evolution of the observable process can be determined from the analysis of the existing condition data. If the observable process can be represented with the exponential distribution, as suggested by [12] and assumed in this article, the appropriate HMM to be used is an Exponential Hidden Markov Models (EHMM).

The use of HMMs to model road deterioration is explored in this article. The article is divided as follows: Section 2 contains a discussion on the relationship between the observed process and the hidden process and an explanation of the EHMM. Section 3 and 4 contain the model formulation and estimation methodology, respectively. Section 5 contains a case study, using information from the national road network of Vietnam. Section 6 contains the conclusions of the work and gives indication of future research needs.

II. Observed and Hidden Processes

A road immediately after it is built is normally considered to be in perfect condition and over time, to deteriorate. The condition of the road can be described in different ways, for example discrete condition states may be used, $i(i=1,\cdots I)$, where condition state 1 may be considered perfect or like perfect and I alarming, i.e. the condition state where it is simply no longer acceptable and an intervention is to be executed. These condition states are often described with attributes, or indicators, e.g. the amount of cracked surface area (m^2). Although these indicators do represent the real condition they are only an indicator of the real condition. The relationship between the values of an indicator and the physical condition is depicted in Fig. 1. In Fig. 1 the values of the indicators increase over time and the physical condition described by condition states is shown.

If a Markov model of deterioration is to be used, transition probabilities need to be estimated, requiring data on the evolution of the values of the condition indicators over time. In many cases, however, this data is insufficient, to accurately estimate the transition probabilities, introducing significant uncertainty into the deterioration model, and therefore in the optimality of the theoretically OIS.

If data on all required condition indicators over time is not available, or can not be collected, it is possible to approximate the transition probabilities of the deterioration process using the data of an single condition indicator over time and limited data used to construct Markov transition probability, exploiting the correlation between the indicator for which data is available and the deterioration process. In this case, a Markov model can be used to model the evolution the condition indicator for which data exists, i.e. the observable process, and the deterioration process, i.e. the hidden process. An illustration of the relationship between the observable and the hidden process is given in Fig. 2, where the transitions between condition states $g(t_r^l=i,\ \ i=1,\cdots,I)$ of the road section l at respective times $t_r^l(r=1,\cdots,T)$ can be deduced from the values of the condition indicator that correspond to each condition state $h(t_r^l)$. Knowing the condition state transitions over time, the transition probabilities π_{ij}, to be used in a Markov model of deterioration can be estimated.

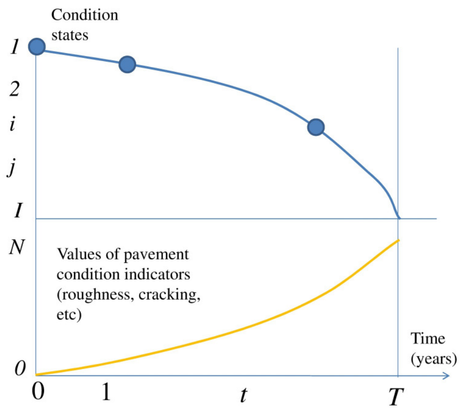

Fig. 1 Relationship between road section condition and a condition indicator

III. The Model

A. Transition probabilities

The model used in this paper to approximate the transition probabilities of the EHMM is an extension of the Maximum-likelihood estimation method proposed by [4]. The explicit mathematical formula for estimating Markov transition probability π_{ij} is given in Eq. (1). A short description of the model's formulation is given in the Appendix.

$$\pi^{ij}(z)=\sum_{k=i}^{j}\prod_{m=i}^{k-1}\frac{\theta_m}{\theta_m-\theta_k}\prod_{m=k}^{j-1}\frac{\theta_m}{\theta_{m+1}-\theta_k}\exp(-\theta_k z), \tag{1}$$

In Eq. (1), the condition states of the road section are described by indexes i,j,k,m, $(i\le k\le m\le j)$. The condition states i and j are the priori and posteriori condition states, respectively, m and k are variables to allow appropriate consideration of condition states between i and j, θ is the deterioration rate, or hazard rate, from respective condition state i , i.e. the rate at which an object goes from condition state i to worse condition states, z is time interval between two inspections. Statistically, the hazard rate θ can be expressed in the multiplicative form $\theta=x\cdot\beta$ (refer to Eq. (23) in the Appendix). Where x is a vector of characteristic variables (traffic volume, change in pavement thickness, temperature, etc) that affect the deterioration process, and β is to be estimated given information on the evolution of condition states and the characteristic variables.

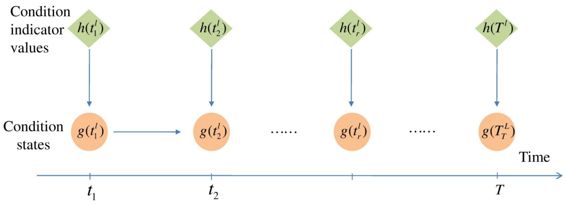

Fig. 2 Illustration of hidden Markov process

B. Exponential Hidden Markov model

In the EHMM, it is assumed that the Markov transition probabilities π_{ij} are used to model the hidden deterioration process, and the available data is the values of a condition indicator. The occurrence of the values of the condition indicator can be represented as the random variable $h(t)$, with t being time.

The probability distribution of $h(t)$ is considered to be dependent on the condition state $g(t)=i(1,\cdots,I)$ at time t. The probability of the condition indicator having a specific value $h(t)=n$, given $g(t)=i$ at time t, is assumed to follow an exponential probability distribution (Eq. (2)),

$$\gamma_{in} = Prob[h(t)=n \mid g(t)=i] = \lambda_i exp(-\lambda_i t), \tag{2}$$

where n is the value of the condition indicator at time t and λ_i is the rate parameter (or parameter of distribution) associated with condition state i. In another words, it is the rate of leaving condition state i.

Eq. (2) is also constrained in that the summation of all possible probabilities γ_{in} must equal to 1 ($\int_{n=1}^{N} \gamma_{in} = 1$), where N is the worst possible value of the condition indicator.
The conditional probability distribution of $h(t)$ with respect to condition state i can then be determined using the joint probability distribution function of condition state and value of condition indicator.

$$Prob[h(t)=n] = Prob[h(t)=n \mid g(t)=i]Prob[g(t)=i] \tag{3}$$
$$= \gamma_{in}\rho_i,$$

where ρ_i is a conditional state probability of the object being in state i. ρ_i can be defined by as $\rho_{i,t} = \rho_{i,t-1}\pi_{ij}$.

Since the values of condition indicators are affected by both deterioration and the improvement resulting from interventions, the likelihood function to be used to determine the values of the condition indicator is:

$$L\{h(t)=n\} = \sum_{i_1=1}^{I}\sum_{i_2=1}^{I}\cdots\sum_{i_T=1}^{I}\rho_{i_1}\gamma_{i_1,n_1}\prod_{t=2}^{T}\pi_{i_{t-1},i_t}\gamma_{i_t,n_t}. \tag{4}$$

Given the exponential distribution representing the relation between the condition states and the value of the condition indicator in Eq. (2), Eq. (4) can be re-written as:

$$L\{h(t)=n\} = \sum_{i_1=1}^{I}\sum_{i_2=1}^{I}\cdots\sum_{i_T=1}^{I}\rho_{i_1} exp(-\lambda_{i_1} 1)\cdot\prod_{t=2}^{T}\pi_{i_{t-1},i_t} exp(-\lambda_{i_t} t) \tag{5}$$

Similar types of expressions for Eqs. (4) and (5) can be found in recent literature [13, 14, 15].

IV. METHODOLOGY

In order to explain the methodology it is convenient to introduce the following terminology. Observed values of the condition indicator for a road section l in a road network with L sections over time are denoted as τ_t^l $(t=1,\cdots,T^l)$, with T^l) as the number of inspections for the road section l. The hazard rate of a road section l is influenced by the changes in the values of characteristic variables, such as traffic volume, thickness of overlay, and weather, etc, which can be referred to Eq. (23) in the Appendix., is referred to as θ^l.

In order to use the observed data it is also necessary to introduce the following two dummies variables:

$$\delta_i^{tl} = \begin{cases} 1 & if \quad g^l(t) = i \\ 0 & otherwise \end{cases} \tag{6}$$

and

$$\delta_{ij}^{tl} = \begin{cases} 1 & if \quad g^l(t-1) = i \quad and \quad g^l(t) = j \\ 0 & otherwise \end{cases} \tag{7}$$

By using these dummy variables and taking the natural logarithm of both sides of the likelihood function in Eq. (5) with respect to the complete set of data, i.e. all information related to the observed and hidden processes, the following complete likelihood function can be obtained:

$$\ln L =$$
$$\sum_{l=1}^{L}\left[\sum_{i=1}^{I}\delta_i^{1l}\ln(\rho_i^l) + \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\sum_{t=1}^{T}\delta_{ij}^{tl}\right)\ln\pi_{ij}^l - \sum_{i=1}^{I}\sum_{t=1}^{T}\delta_i^{tl}\lambda_i^{t,l}n_t\right]. \tag{8}$$

The values of the parameters in the set $\Theta = (\rho,\pi,\lambda)$ are then estimated by determining the optimal solution to Eq. (8), i.e. the most likely values of the hidden process knowing the values of the observable process.

This is done using the Maximum-likelihood estimation (MLE) method, a popular statistical method used to fit statistical models to observed data, and the Baum-Welch algorithm, a particular case of expectation-maximum (EM) algorithm [16, 17], which is suitable for determining optimal solutions for hidden Markov models when there is incomplete data. Using the Baum-Welch algorithm the likelihood in Eq. (8) can be simplified as:

$$L_c = Prob\{g(t)=i \mid h(t)=n\}\cdot Prob\{h(t)=n\} \tag{9}$$

The maximization of Eq. (9) is done by performing the EM algorithm (A detailed explanation for the M-step and E-step in the EM algorithm can be found in [16, 18]). The steps of E-M algorithm is outlined in Table I.

TABLE I
OUTLINE OF EM-ALGORITHM

1. Set the initial values for $\Theta = (\rho, \pi, \lambda)$
2. Start the LOOP,
3. Do the E-step to estimate the values of the dummy variables δ_i^τ and δ_{ij}^τ given the initial values of Θ,
4. Do the M-step to compute the values of Θ by maximizing the likelihood function Eq. (9)
5. Check the convergent condition for the values of Θ,
   a. if not satisfied, the iteration goes back to step 2,
   b. if satisfied, stop the iteration

V. EXAMPLE

A. Problem

The feasibility and usefulness of the proposed EHMM is demonstrated in this section using inspection data on 6'510 asphalt concrete road sections of 1 kmm in length in Vietnam. Two nationalwide inspections were carried out in 2001 and 2004. A significant amount of data on roughness was recorded, while only a relatively small amount of data on other, percentage of surface cracking, an number of potholes was recorded (Table II). As no composite index was used to indicate the condition of the pavement, we define a composite condition index (CCS). The value of the CCS is estimated using the weighted values of the international roughness index (IRI) and the percentage of the surface area with horizontal, longitudina or alligator crack, as shown in:

$$CCS = w_r \cdot \frac{r}{16} + w_c \cdot \frac{c}{100}, \tag{10}$$

where
r - the international roughness index (mm/km)
c - the surface area with horizontal, longitudinal or alligator cracks ($\%$)
w - weighting factors, where $w_r + w_c = 1$.

In Eq. (10), the denominators 16 (mm/m) and 100 (%) are the ultimate values of the IRI and the percentage of surface area that is cracked that can be measured.

The CCS used in this work, is introduced to obtain a superior approximation of the true physical condition of road section than that provided alone by the IRI. The building of the CCS on the IRI and the percentage of cracked surface area of the pavement is strongly supported by the empirical studies, where it is been demonstrated that there is weak correlation between the evolution of the IRI and the percentage of cracked surface area [19]. The condition of the road section was defined using the 5 discrete condition states shown in Table III.

TABLE II
OVERVIEW OF DATA

Condition indicator	Measurement	Unit	Number of data	
Roughness	International roughness index	mm/m	6'510	6'510
Cracking of surface area	Horizontal, longitudinal, and alligator crack	%	1'237	1'237
Potholes		numbers	No	No

TABLE III
NOTATIONS OF CONDITION STATES

Condition states	Equivalent CCS	Remark
1	[0-0.2]	Very good
2	(0.2-0.4]	Good
3	(0.4-0.6]	Fair
4	(0.6-0.8]	Poor
5	0.8	Very poor

B. Model

1)Transition probabilities using the multi-stage exponential Markov hazard model: The transition probabilities of the Markov model used to model the evolution of the CCS were first estimated using the 1'237 records where data on both roughness and surface cracking was available (Table I). Using the model proposed by [4] and the traffic volume and the pavement thickness as covariates, i.e. two significant factors influencing the deterioration process, the values of the β parameter of the multi-stage exponential Markov model, were determined (Table IV). The values of the hazard rates θ_i for each condition state i were estimated using Eq. (23) in the Appendix. Statistically, the values of the unknown parameter β and their corresponding *t-values* infer that traffic volume has more impact on the deterioration speed than surface thickness, where a road section is in condition state 2. Once a road section is in condition state 3 and 4, however, surface thickness has more impact on the deterioration speed than traffic volume.

TABLE IV
VALUES OF THE PARAMETERS β IN THE MULTI-STAGE EXPONENTIAL MARKOV MODEL

Condition states	Absolute $\beta_{i,1}$	Thicknes s $\beta_{i,2}$	Traffic volume $\beta_{i,3}$	Hazard rate θ_i
1	0.305 (26.027)	-	-	0.305
2	0.179 (2.655)	0.499 (2.443)	1.057 (2.417)	0.489
3	-	2.646 (9.629)	-	0.894
4	-	1.808 (6.542)	-	0.611

Using Eq. (1) and the data collected for the 1'237 road sections with data on the roughness and cracking, the expected Markov transition probabilities were obtained (Table V). In addition, the distribution of condition states over time is drawn in Fig. 3.

TABLE V
MARKOV TRANSITION PROBABILITIES ESTIMATED USING THE MULTI-STAGE EXPONENTIAL MARKOV MODEL

Condition states	1	2	3	4	5
1	0.737	0.205	0.043	0.013	0.002
2	0	0.613	0.247	0.113	0.027
3	0	0	0.409	0.423	0.168
4	0	0	0	0.543	0.457
5	0	0	0	0	1

International Journal of Architectural, Civil and Construction Sciences
ISSN: 2415-1734
Vol:6, No:1, 2012

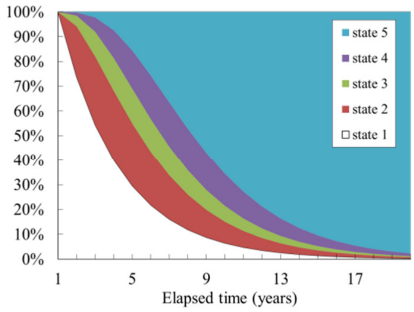

Fig. 3 Distribution of condition states on 1'237 road sections using the multi-stage exponential Markov model

2) *Initial parameter values for the exponential hidden Markov model (EHMM)*: The transition probabilities of the Markov model of the hidden process were determined using the EHMM model and the roughness and cracking data for all 6'510 road sections. The initial condition state probabilities of all road sections were assumed to be ρ=(1,0,0,0,0)., i.e. all road sections were in a like new condition at the time of construction. The hazard rates and Markov transition probabilities determined using the data from 1'237 road sections were used as initial values.

In order to use the transition probabilities of the observable process to model the hidden process, i.e. in the EHMM model, it was necessary to determine the initial values of the exponential rate parameter λ_i for each condition state *i* of the hidden process. These values are estimated by mapping the roughness measurement to each condition state used to describe the deterioration process, using the MLE approach (Eq. (11)). It is noted that the use of prior knowledge in the form of expert opinions can also be used to determine λ_i. Under such case, the values of λ_i are considered as constants. This approach was not used in this study.

$$\overline{\eta}_i = \frac{1}{n}\sum_{e=1}^{n}\eta_{i,e} \tag{11}$$

Where, η_i is the observed value and $\overline{\eta}_i$ is its mean.

The rate parameter $\overline{\lambda}_i$ can then be defined as:

$$\overline{\lambda}_i = \frac{1}{\overline{\eta}_i} \tag{12}$$

For the specific problem investigated , the values of the rate parameters were determined to be $\overline{\lambda}_i$=(0.500, 0.200, 0.143, 0.111, 0.077).

C. Results of the EHMM

Using the EHMM model, programed in R-language, and the initial values of the parameters of the model, the new values of the hazard rate were determined θ_i=(0.287, 0.344, 0.841, 0.551)., i.e. the speed at which the CCS leaves each of the defined condition state. Substituting these new values into Eq. (1), the new Markov transition probabilities are obtained (Table VI). The distribution of residuals of our estimation, i.e. the value representing unexplained variation after fitting the data to regression model or in other words the difference between the observed value and the value determined using the regression model has a normal distribution, with a nice bell shape (Fig. 4). And their sum-around the mean is 0, indicating that the regression model is appropriate for the data [20].

TABLE VI
MARKOV TRANSITION PROBABILITIES ESTIMATED USING THE EHMM BASED ON THE INFORMATION FROM ALL 6'510 ROAD SECTIONS

Condition states	1	2	3	4	5
1	0.745	0.211	0.033	0.009	0.002
2	0	0.709	0.192	0.082	0.017
3	0	0	0.431	0.421	0.148
4	0	0	0	0.576	0.424
5	0	0	0	0	1

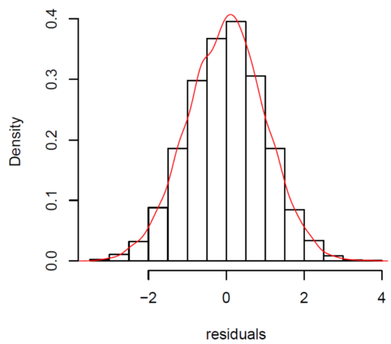

Fig. 4 Distribution of residuals

The expected distribution of condition states over time, i.e. the mean value determined from the EHMM for all 6'510 road sections is almost linear from condition state 1 to 3, with almost equal amounts of time spent in each condition state 1 and 2. However, there is a sharp decrease of pavement quality when its condition reaches to condition state 3. The average time for the road section to stay in condition state 3 and 4 is about 2 years. As a result, it takes on average 9.5 years for a typical road section to go from new to condition state 5 (Fig. 5). This speed of deterioration is fast if compared to asphalt concrete road sections in developed countries like Japan or Switzerland where similar deterioration might be expected that take 25 years, but it more or less what is expected in Vietnam. This fast deterioration is most likely attributable to the lower quality of construction of the road sections, combined with the heavy annual traffic volume and prevalent soft sub-soil conditions [21].

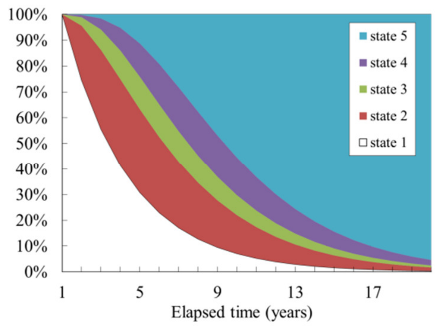

Fig. 5 Distribution of condition states on 6'510 road section using the EHMM model

In addition to the estimation of model on 6'510 data of road sections, we create another database consisting of 5'273 road sections (6'510-1'237), where having only information of roughness, while there is no information on cracking. The mean value of rate parameter is $\overline{\lambda}_i$ =(0.312, 0.164, 0.121, 0.091, 0.067), which results in the hazard rate θ_i =(0.288, 0.355, 0.845, 0.586). The Markov transition probabilities matrix for the case of 5'273 road sections is shown in Table VII.

TABLE VII
MARKOV TRANSITION PROBABILITIES ESTIMATED USING THE EHMM BASED ON THE INFORMATION FROM ALL 5'273 ROAD SECTIONS

Condition states	1	2	3	4	5
1	0.750	0.209	0.031	0.009	0.001
2	0	0.701	0.197	0.083	0.019
3	0	0	0.430	0.414	0.156
4	0	0	0	0.557	0.443
5	0	0	0	0	1

Comparing the condition state distributions estimated using the multi-stage exponential Markov model on the 1'237 road sections and the EHMM model on the 5'273 and 6'510 road sections (Fig. 3, Fig. 6, and Fig. 5, respectively), it can be see that the deterioration predicted using the EHMM model is slightly longer than that predicted using the the multi-stage exponential hazard model. The two models also predict different distribution of condition states over time. In addition, deterioration curves representing estimation results of model on three databases are shown in Fig. 7, Where the EHMM predicts a faster initial deterioration and slower final deterioration than the multi-stage exponential hazard model. Although it is believed that the result produced using the EHMM is more accurate than that produced using the multi-stage exponential hazard model more research, on a large data set is required to prove it.

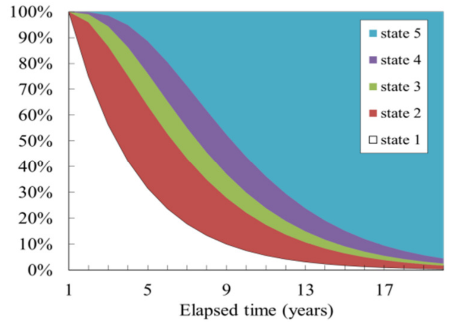

Fig. 6 Distribution of condition states on 5'273 road section using the EHMM model

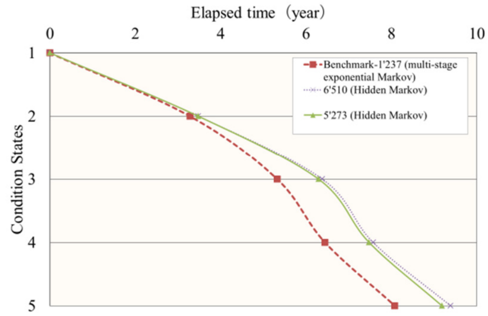

Fig. 7 Performance curves

VI. CONCLUSION

In this paper the potential use of the exponential hidden Markov model to predict the deterioration of road sections when incomplete inspection data is available is demonstrated. It was demonstrated for the case where the values of an overall performance indicator, built from two indicators, is required to determine the optimal intervention strategy, but only the complete values of one of the indicators is complete. Only an incomplete set is available for the other performance indicator. The evolution of the overall performance indicator is described as an exponential distribution, and is considered to be hidden. The hidden Markov model was then used to estimate the evolution over time of the overall performance indicator from the incomplete data. Data collected on the Vietnamese national road network was used in the case study.

APPENDIX

MATHEMATICAL FORMULATION OF THE MULTI-STAGE EXPONENTIAL MARKOV MODEL

From condition data from two temporally consecutive inspections t and $t+1$, the Markov transition probability can be described as follows:

$$Prob[g(t+1)=j \mid g(t)=i]=\pi^{ij}. \tag{13}$$

Markov transition probability matrix can be written in the following form:

$$\Pi=\begin{pmatrix} \pi^{11} & \cdots & \pi^{1I} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \pi^{II} \end{pmatrix}, \tag{14}$$

where

$$\pi^{ij} \geq 0\,(i,j=1,\cdots,I)$$
$$\pi^{ij}=0\,(\quad when \quad i>j)$$
$$\sum_{j=i}^{I}\pi^{ij}=1 \tag{15}$$

The time that an object is expected to stay in condition state i is assumed to be a stochastic variable, with the probability density function $f_i(\zeta_i)$ and distribution function $F_i(\zeta_i)$. The conditional probability, to which condition state i at time y_i reaches condition state $i+1$ at $y_i+\Delta_i$, can be expressed as hazard function $\lambda_i(y_i)\Delta y_i$:

$$\lambda_i(y_i)\Delta y_i = \frac{f_i(y_i)\Delta y_i}{\tilde{F}_i(y_i)}, \tag{16}$$

where $\tilde{F}_i(y_i) = 1 - F_i(y_i)$ is referred as the survival function of the object in condition state i during the time interval from $y_i = 0$ to y_i.

Since it is assumed that the deterioration process satisfies the Markov property and the hazard function is independent of the time instance y_i the hazard rate is constant and positive:

$$\lambda_i(y_i) = \theta_i. \tag{17}$$

The time that an object is in condition state i longer than the time instance y_i, is referred as the value of the survival function $\tilde{F}_i(y_i)$ and can be expressed in the exponential form as:

$$\tilde{F}_i(y_i) = \exp(-\theta_i y_i). \tag{18}$$

When the y_i equals z of the inspection period z_i between $[t, t+1)$, the value of the survival function is identical to the transition probability π^{ii}, therefore:

$$\tilde{F}_i(t+z \mid \zeta_i \geq t) = Prob\{\zeta_i \geq t+z \mid \zeta_i \geq t\} \tag{19}$$

$$= \frac{\exp\{-\theta_i(t+z)\}}{\exp(-\theta_i t)} = \exp(-\theta_i z),$$

$$Prob[g(t+1) = i \mid g(t) = i] = \exp(-\theta_i z). \tag{20}$$

By defining the subsequent conditional probability of condition state j to i, with respect to the actual interval time z of inspection, a general mathematical formula for estimating the Markov transition probability can be defined:

$$\pi^{ij}(z) = Prob[h(t+1) = j \mid h(t) = i] \tag{21}$$

$$= \sum_{k=i}^{j} \prod_{m=i,\neq k}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} \exp(-\theta_k z),$$

where

$$\prod_{m=i,\neq k}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} \exp(-\theta_k z) \tag{21-a}$$

$$= \prod_{m=i}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} \prod_{m=k}^{j-1} \frac{\theta_m}{\theta_{m+1} - \theta_k} \exp(-\theta_k z),$$

and

$$\begin{cases} \prod_{m=i}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} = 1 & (k = i) \\ \prod_{m=k}^{j-1} \frac{\theta_m}{\theta_{m+1} - \theta_k} = 1 & (k = j) \end{cases} \tag{21-b}$$

$$(i = 1, \cdots, I-1; j = i+1, \cdots, I).$$

Transition probability from condition state i to absorbing condition state I is eventually defined in the following equation:

$$\pi^{iI}(z) = 1 - \sum_{j=i}^{I-1} \pi^{ij}(z) \; (i = 1, \cdots, I-1). \tag{22}$$

The likelihood function of hazard rate θ_i can be expressed in multiplicative form with characteristic variable x and unknown parameter β'_i.

$$\theta_i = \theta_i(x) = x\beta'_i. \tag{23}$$

The remaining time of the object in condition state i $RMD_i(x)$ is then given by the survival probability of condition state i over continuous time.

$$RMD_i(x) = \int_0^\infty \tilde{F}_i(y_i \mid \theta_i(x)) dy_i \tag{24}$$

$$= \int_0^\infty \exp\{-\theta_i(x) y_i\} dy_i = \frac{1}{x\beta'_i}.$$

The average time of the object in condition state $j(>1)$ can be defined by the summation of the times over the range of condition states counted from $i = 1$:

$$ET_j(x) = \sum_{i=1}^{j} \frac{1}{x\beta'_i}, \tag{25}$$

where ET_j stands for average time of the object being in condition state j.

References

[1] K. Kaito, K. Kazuhiko, H. Hayashi, and K. Kobayashi, *"Forecasting cracking process of the road by using the step-wide hazard model", JSCE Journal (in Japanese),* vol. 63, p386-402, 2007.

[2] Z.S. Nakat, and S. Madanat, "Stochastic Duration Modeling of Pavement Overlay Crack Initiation", *Journal of Infrastructure Systems,* vol. 14(3), p185-192, 2008.

[3] H.C. Shin, and S. Madanat, *"Development of stochastic model of pavement distress initation"*, *Journal of Infrastructure Planning and Management*, vol. IV-61(744), p61-67, 2003.

[4] Y. Tsuda, K. Kaito, K. Aoki, and K. Kobayashi, "Estimating Markovian transition probabilities for bridge deterioration forecasting", *Journal of Structural Engineering and Earthquake Engineering,* vol. 23(2): 241-256, 2006.

[5] K. Golabi, and R. Shepard, *Pontis:* "A System for Maintenance Optimization and Improvement of US Bridge Networks", *Interfaces,* vol. 27(1),p71-88, 1997.

[6] R. Guido, H. Rade, B.T. Adey, and B. Eugen, "Condition evolution in bridge management systems and corrosion-included deterioration", *Journal of Bridge Engineering,* vol. 9(3)p268-277, 2004.

[7] S.K. Sinha, and A.K. Mark, "Intelligent System for Condition Monitoring of Underground Pipelines", *Computer-Aided Civil and Infrastructure Engineering*, vol. 19(1), p42-53, 2004.

[8] S.K. Sinha, and A.M. Robert, "Probabilistic based integrated pipeline management system", Tunnel*ing and Underground Space Technology,* vol. 22(5-6): 543-552, 2007.

[9] M.T. Shahin, *Pavement Management for Airports, Roads, and Parking Lots*, Springer, 2005.

[10] N. Lethanh, 2009. "*Stochastic Optimization Methods for Infrastructure Management with Incomplete Monitoring Data"*, PhD dissertation, Graduate School of Engineering, Kyoto University, September, 2009.

[11] Y. Ephraim, and N. Merhav, "Hidden Markov processes", *Information Theory, IEEE Transactions*, vol. 48(6),p1518-1569, 2002.

[12] W.D.O. Paterson, "International roughness index: relationship to other measures of roughness and riding quality", *Transportation Research Record*, vol. 1084, p49-59, 1990.

[13] W. Zucchini, and L.M. Lain, *Hidden Markov Models for Time Series*, Taylor and Francis. 2009.

[14] O. Capper, E. Moulines, and T. Ryden. 2005. *Inference in Hidden Markov Models.* Springer, 2005.

[15] R. Paroli, R. Giovanna, and S. Luigi, "Poisson Hidden Markov Models for Time Series of Overdispersed Insurance Counts", *Causality Actuarial Society*,: 461-472, 2000.

[16] G.J. McLachlan, and T. Krishnan, *The EM Algorithm and Extensions*. New York. Wiley, 1997.

[17] L.R. Welch, "Hidden Markov Models and the Baum-Welch Algorithm", *IEEE Information Theory Society Newsletter*, vol. 53(4), p10-13,2003.
[18] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum likelihood from incomplete data via the EM Algorithm", *Journal of the Royal Statistical Society Series B*, vol. 39, No. 1, p1-38, 1977.
[19] A.H. Lisa, E. Jackson, C.E. Dougan, and S. Choi. "Models Relating Pavement Quality Measures", *Transportation Research Record: Journal of the Transportation Research Board,* vol. 1869: p119-125, 2004.
[20] D.C. Montgomery, and G.C. Runger, *Applied Statistics and Probability for Engineer*, With CD. Phoenix Usa. Wiley India Pvt. Ltd, 2007.
[21] N. Lethanh, T. Nguyendinh, K. Kaito, and K. Kobayashi, "A Benchmarking Approach to Pavement Management: Lessons from Vietnam", *Infrastructure Planning Review,* vol. 26(1), p101-112, 2009.