LOCAL MIXTURE HAZARD MODEL FOR BENCHMARKING THE DETERIORATION PROCESS AMONG INFRASTRUCTURE TECHNOLOGIES

by Le Thanh NAM **, Kiyoyuki KAITO*** and Kiyoshi KOBAYASHI****

1. Introduction

In an infrastructure system comprising of many groups categorized by technological differences, the overal degradation of the entire system is the average value, which largely depends on the deterioration rank of individual group. Each group is regarded as one type of technology. The deterioration of individual group is subjected to be different from the others due to the rich variety in design, construction, environment and operation condition. In regard to the selection of technology based upon the best actual performance, it is necessary to study the deterioration behavior of each group by mean of hetegoneity factor. Only when the heterogeneity factor is emperically determined, had the better choice on optimal technology would become feasible. This paper focuses upon the benchmarking model using local mixing mechanism in order to define the value of hetegoneity factor and further describe and compare the deterioration process of each group of technology.

In regard to the deterioration forecasting model, in recent years, the use of Markov chain based analysis has been one of the major innovations. Hazard model helps users to predict hazard rates, life spans and deterioration curves of infrastructure given the historical inspected condition states and other variables concerning various environment impacts. The application of the Markov chain model has gained its high recognition for its flexibility of modeling and high operability. In the Markov chain model, the condition states of infrastructure component are ranked in some discrete states. The values of condition states are recorded through actual inspection over periods of time. And thus, a simulation of deterioration is understood to follow the transit of condition states along with duration[1-2].

The proposed model in this research discusses the integration of hetegeneity factor into conventional markovian hazard model that as developed by Tsuda et at[3]. The hetegoneity factor is followed by local mixing mechanism. Estimation approach enables us to determine not only the deterioration of the entire group of infrastructure but also the individual. An empirical analysis on the Vietnamese highway system is carried out to illustrate the application of local mixture hazard model in the real world.

2. Benchmarking Hazard Model

(1) The Deterioration Process and Markovian Hazard Model

The deterioration process of infrastructure under time-homogeneous Markov chain process is defined on state space $S=\{1,\dots,I\}$. Here, the rating $i(1,\dots,I)$ reflexs the healthy status of road sections with $i=1$ to be at its healthiest and $i=I$ to be its worst. The probability, at which the condition state observed at time t expresses as $w(t)=i$ change to $w(t+1)=j$, can be described as conditional probability in the following equation

$$Prob[h(t+1)=j\,|\,h(t=i]=\pi_{ij} \tag{1}$$

The transition probabilities are expressed by a matrix of dimension *(i,j)*. Here, it is understood that different time intervals give different transition probabilities. The highest level when *i=I* is called absorbing state. Using the database of past inspections, the transition probability can be estimated. In this section, we only give an outline of the estimation method for ease to readers.

The life expectancy of a condition state i is assumed to be a stochastic variable with the probability density function $f_i(\zeta_i)$ and the distribution function $F_i(\zeta_i)$ The conditional probability that the condition state i at time y_i of a component reaches to

*Keywords: Markov hazard model, optimal inspection and repair, pipeline system

** Member of JSCE, Graduate Student., Dept of Urban Management., Kyoto University,
(Katsura, Kyotoshi, Japan TEL 075-383-3224, FAX 075-383-3224, E-mail: info.tna@t06.mbox.media.kyoto-u.ac.jp)
*** Member of JSCE, Associate Professor., Graduate School of Engineering., Osaka University,
(Yamada-oka, Suita, Osaka, Japan TEL 066-879-7589, FAX 066-879-7601, E-mail: kaito@ga.eng.osaka-u.ac.jp)
**** Member of JSCE, Professor., Dept of Urban Management., Kyoto University,
(Katsura, Kyotoshi , Japan TEL 075-383-3222, FAX 075-383-3224, E-mail: kkoba@psa.mbox.media.kyoto-u.ac.jp)

condition state $i+1$ at $y_i + \Delta_i$ can be expressed in the following hazard function

$$\lambda_i(y_i)\Delta y_i = \frac{f_i(y_i)\Delta y_i}{\tilde{F}_i(y_i)} \tag{2}$$

where $\tilde{F}_i(y_i) = 1 - F_i(y_i)$ is referred as the survival function of a transition in the condition state i during the time interval $y_i = 0$ to y_i. It is assumed that the deterioration process of a component satisfied the Markov property and the hazard function is independent of the time instance y_i on the sample time-axis. Thus, for a fixed value of $\lambda_i > 0$, it can be defined as $\lambda_i(y_i) = \lambda_i$ The value of survival function $\tilde{F}_i(y_i)$ can be obtained by using exponential hazard function $\tilde{F}_i(y_i) = \exp(-\lambda_i y_i)$. Considering the possible inspection time, y_A for instance, the condition state are observed as i at time y_A and keep remaining constant at time $y_A + z_i, (z \geq 0)$. The conditional probability for this event to happen can be defined as

$$\tilde{F}_i(y_A + z \mid \zeta_i \geq y_A) = Prob\{\zeta_i \geq y_A + z \mid \zeta_i \geq y_A\} = \frac{\exp\{-\lambda_i(y_A + z)\}}{\exp(-\lambda_i y_A)} = \exp(-\lambda_i z) = \pi_{ii}(z) \tag{3}$$

where, z indicates the interval between the two inspection times. Equation (3) expresses the probability $Prob[h(y_B) = i \mid h(y_A) = i]$ as the Markov transition probability π_{ii}. Equation (3) also illustrates the hazard rate λ_i and the time interval z are the only two parameter that are required for the calculation of the transition probability π_{ii}.

By defining the subsequent conditional probability of the condition state *i* to *j*, with respect to the actual interval time *z* of inspection, general form for transition probability is formulated in equation (4).

$$\pi_{ij}(z) = Prob[h(y_B) = j \mid h(y_A) = i] \quad = \sum_{l=i}^{j} \prod_{m=i}^{l-1} \frac{\lambda_m}{\lambda_m - \lambda_l} \prod_{m=l}^{j-1} \frac{\lambda_m}{\lambda_{m+1} - \lambda_l} \exp(-\lambda_l z) \quad (i = 1,\cdots,I-1;\, j = i+1,\cdots,I) \tag{4}$$

For convenience of mathematical manipulation, we define

$$\prod_{m=i,\neq l}^{j-1} \frac{\lambda_m}{\lambda_m - \lambda_l} \exp(-\lambda_l z) = \prod_{m=i}^{l-1} \frac{\lambda_m}{\lambda_m - \lambda_l} \prod_{m=l}^{j-1} \frac{\lambda_m}{\lambda_{m+1} - \lambda_l} \exp(-\lambda_l z) \tag{5}$$

(2) Local Mixing Mechanism of Hazard Model

It has been realized that similar group of individual infrastructure components can be exerted to have different deterioration speeds. To express these differences, the term "heterogeneity factor" is employed. Here, the letter ε^k denotes the heterogeneity parameter, which infers the change of characteristic of a peculiar hazard rate to an infrastructure component $k\,(k = 1,\cdots,K)$. Thus, the mixture index hazard function (2) can be expressed as

$$\lambda_i^k = \tilde{\lambda}_i^k \varepsilon^k \; (i = 1,\cdots,I-1;\, k = 1,\cdots,K) \tag{6}$$

The value of ε^k is always greater than 0. Importantly, the deterioration speed of component k is fast when value of ε^k increased in comparison with the rate of standard hazard $\tilde{\lambda}_i^k$. It is also noted from (10) that the same random variable ε^k is included in the mixture index hazard function of all ratings. ε^k is understood to be in a form of function or stochastically distribution. The Markov transition probability $\pi_{ij}^k(z^k : \bar{\varepsilon}^k)$ that the rating changes to $j\,(> i)$ between y_A^k and $y_B^k = y_A^k + z^k$ will be formed from (7) and (10) as

$$\pi_{ij}^k(z^k : \bar{\varepsilon}^k) = \sum_{l=i}^{j} \prod_{m=i,\neq l}^{j-1} \frac{\tilde{\lambda}_m^k}{\tilde{\lambda}_m^k - \tilde{\lambda}_l^k} \exp(-\tilde{\lambda}_l^k \varepsilon^k z^k) \quad = \sum_{l=i}^{j} \psi_{ij}^l(\lambda^k) \exp(-\tilde{\lambda}_l^k \varepsilon^k z^k) \qquad k = 1,\cdots,K) \tag{7}$$

where,

$$\psi_{ij}^l(\lambda^k) = \prod_{m=i,\neq l}^{j-1} \frac{\tilde{\lambda}_m^k}{\tilde{\lambda}_m^k - \tilde{\lambda}_l^k} \tag{8}$$

Equation (7) expresses the dependence of transition probabily only on the function of average hazard rate.

Regarding the heterogeneity factor ε^k, it is not easy to infer ε^k to follow a particular function. Several investigations were on assuming function of ε^k as stochastic distribution such as: Gamma, Lognormal distributions etc. However, the results are not always satisfied in fitting with actual transition. In order to cope with this obstacle, one possible rule is to consider the form taken by mixture distribution as when ε^k has little dispersion so that the departure from homogeneity is small [4]. To express the relation of transition probability and hetegeroneity, equation (6) can be further described in the following form.

$$\tilde{\pi}_{ij}(z) = \int_0^{\infty} \pi_{ij}(z : \varepsilon) f(\varepsilon) d\varepsilon \quad (i = 1,\cdots,I-1) \tag{9}$$

For ease of mathematical expression, let assume the local mixture transition probability as the exponential function $f_{mix}(\varepsilon, z, \lambda)$ with mix as indication for local mixture and define the following equation.

$$f_{mix}(\varepsilon, z, \lambda) = \int f(\varepsilon, z, \lambda) dH(\varepsilon) \tag{10}$$

where, $dH(\varepsilon)$ is arbitrary distribution around the mean of 1. $f(\varepsilon, z, \lambda)$ is exponential family, and thus, can be further expressed as $f(\varepsilon,z,\lambda)=exp(-\varepsilon\lambda z)$. The expected value of $f(\varepsilon,z,\lambda)$ with respect to ε is likely a function of ε about its mean, and can be taken to be unity with no loss of generality as long as the mean exits in the following form.

$$exp(-\varepsilon\lambda z) = e^{-\lambda z}(1+(\varepsilon-1)(-\lambda z)+\frac{(\varepsilon-1)^2}{2!}(-\lambda z)^2+... \tag{11}$$

Without lossing generality, the expectation of this series can be expressed in quadratic form. This assumption is proved to satisfy the exponential family as it turn to produce very attractive statistical property. Equation (11) becomes

$$E(e^{-\varepsilon\lambda z}) \approx e^{-\lambda z}\{1+\frac{\sigma^2(\lambda z)^2}{2}\} \tag{12}$$

From this approximation approach, equation (7) and (9) can be rewritten in equation (13)

$$\tilde{\pi}_{ij}(z) = \sum_{l=i}^{j}\psi_{ij}^{l}(\tilde{\lambda})e^{-\tilde{\lambda}_l z}\{1+\frac{\sigma^2(\tilde{\lambda}_l z)^2}{2!}\} \tag{13}$$

The deterioration process of sample k can be expressed by using mixture index hazard function $\lambda_i^k(y_i^k) = \tilde{\lambda}_i^k\varepsilon^k$. The hazard rate $\tilde{\lambda}_i^k$ depends on the characteristic vector of road component and suppose to change to the vector x^k as follows $\tilde{\lambda}_i^k = x^k\beta_i'$. Where $\beta_i = (\beta_{i,1},\cdots,\beta_{i,M})$ is a row vector of unknown parameters and the symbol $'$ indicates the vector is transposed. From equation (13), the standard hazard rate in each rating can be expressed by the mean of the probability distribution of hazard rate $\tilde{\lambda}_i^k$ and the heterogeneity ε^k. The average Markov transition probability is expressible by (13) when using row vector $\bar{x}^k$ of the infrastructure component ($\bar{x}^k = (\bar{x}_1^k,\cdots,\bar{x}_M^k)$ indicating the observed value of variable m for the sample k). In addition, the transition probability also depends on the inspection time interval $\bar{z}^k$ when data is observed. Thus, it is $\tilde{\pi}_{ij}^k(\bar{z}^k,\bar{x}^k:\theta)$ with $(\bar{z}^k,\bar{x}^k)$ and $\theta = (\beta_1,\cdots,\beta_{I-1},\sigma)$ for the average Markov transition probability $\tilde{\pi}_{ij}^k$.

$$L(\theta,\Xi) = \prod_{i=1}^{I-1}\prod_{j=i}^{I}\prod_{k=1}^{K}\left\{\tilde{\pi}_{ij}^k(\bar{z}^k,\bar{x}^k:\theta)\right\}^{\bar{\delta}_{ij}^k} \tag{13}$$

where, $\delta^k = (\bar{\delta}_{11}^k,\cdots,\bar{\delta}_{I-1,I}^k)$ is a dummy variable vector and takes value *1* at $h(\bar{t}^k) = i, h(\bar{y}^k) = j$ and *0* otherwise. Since $\theta = (\beta,\sigma)$, and $\tilde{\pi}_{ij}^k(\bar{z}^k,\bar{x}^k:\theta)$ is a rating transition probability at the initial time. It can be expressed as

$$\tilde{\pi}_{ij}(\bar{z}^k,\bar{x}^k:\theta) = \sum_{l=i}^{j}\psi_{ij}^l(\lambda)e^{-\bar{x}^k\beta_l'\bar{z}^k}\{1+\frac{(\sigma\bar{x}^k\beta_l'\bar{z}^k)^2}{2!}\} \quad (i=1,\cdots,I-1;\ j=i+1,\cdots,I) \tag{14}$$

where,

$$\psi_{ij}^l(\lambda^k) = \prod_{m=i,\neq l}^{j-1}\frac{\bar{x}^k\beta_m'}{\bar{x}^k\beta_m'-\bar{x}^k\beta_l'} \tag{15}$$

Since $\bar{\delta}_i^k, \bar{z}^k, \bar{x}^k$ are known from the inspection, the likelihood functions are functions of β, μ. In the method of maximum likelihood, $\hat{\theta} = (\hat{\beta},\hat{\sigma})$ that maximizes (13) will be presumed. Functions (13) can be defined as the log-likelihood function as follow

$$\ln L(\theta,\Xi) = \sum_{i=1}^{I-1}\sum_{j=1}^{I}\sum_{k=1}^{K}\bar{\delta}_{ij}^k\tilde{\pi}_{ij}^k(\bar{z}^k,\bar{x}^k:\theta) \tag{16}$$

The estimation of $\hat{\theta} = (\hat{\theta}_1,\cdots,\hat{\theta}_{(I-1)M})$ can be obtained by solving the optimality conditions $\frac{\partial \ln L(\theta,\Xi)}{\partial\theta_i} = 0$

When the value of $\hat{\theta}$ is obtained, from equation (10), the following likelihood function is applied to estimate value of ε^k

$$\ln\ell^k(\xi^k:\hat{\theta},\varepsilon^k) = \ln f(\varepsilon^k:\hat{\sigma}) + \sum_{i=1}^{I-1}\sum_{j=i}^{I}\bar{\delta}_{ij}^k\pi_{ij}^k(\bar{z}^k,\bar{x}^k:\hat{\beta},\varepsilon^k) \quad (k=1,\cdots,K) \tag{17}$$

The optimum value of ε^k can be obtained by solving optimality condition $\frac{\partial\ln\ell^k(\xi^k:\hat{\theta},\varepsilon^k)}{\partial\varepsilon^k} = 0$

3. Empirical Analysis

Two set of visual inspection data on the Vietnamese national highway system in the year 2001 and 2004. In which, 1237 highway sections were selected. We further categorized 1237 sections into 8 groups according to their similarity in term of the traffic volume and the asphalt overlay thickness. The range of rating is also defined for the Markov model since there has been no national standard of rating for highway in Vietnam[5]. We selected a range of rating in the space of *S=(1,...,5)*.

In the first step, we estimated the average transition probabilities of the highway system by using hazard model that is explained in (7). The two parameter used for estimation are the thickness of overlay asphalt (x_2) and the traffic volume (x_3), which are subjected to change over two time point of inspection. In general, the general form of the hazard function is explainable by $\lambda_i^k = \beta_{i,1} + \beta_{i,2} x_2^k + \beta_{i,3} x_3^k$ with *(i=1,...,4; n=1,...,N)*, with β, N indicating unknown parameters and the number of samples respectively. Table III shows the result of the maximum likelihood estimations, in which, the value of unknow parameters $\hat{\beta}$ are obtained with the respective *t*-values of each explanatory variable.

Table 1: Exponential hazard model results

State	Absolute $\beta_{i,1}$	Surface Thickness $\beta_{i,2}$	Traffic Volume $\beta_{i,3}$
1	0.3052 (26.027)	-	-
2	0.1792 (2.6557)	0.4996 (2.4433)	1.0570 (2.4174)
3	-	2.6461 (9.6294)	-
4	-	1.8089 (6.5421)	-

Note) *t*-values are shown in parenthesis

Table 2: Average transition probability matrix

State	1	2	3	4	5
1	0.5431	0.2773	0.0998	0.0574	0.0224
2	0	0.3756	0.2522	0.2348	0.1374
3	0	0	0.1673	0.4022	0.4305
4	0	0	0	0.2945	0.7055
5	0	0	0	0	1

Table IV shows the average transition probability when taking the whole set of data into calculation of the exponential hazard model. The values are coressponding to the mean of 1 for the heteroneity factor. Benchmark deterioration curve is drawn as a result of this transition probability.

In the second step, the local mixture model is applied to estimate the heterogeneity factor ε^k. In this application, the number of groups is 8. The values of ε^k are estimated to be (0.6495, 0.9909, 0.9957, 0.5109, 1.1931, 0.5665, 0.5112, 1.1545) respectively to particular group. Thus, a combination of step 1 and step 2 makes it possible to draw the deterioration curves respectively in "Fig. 1".

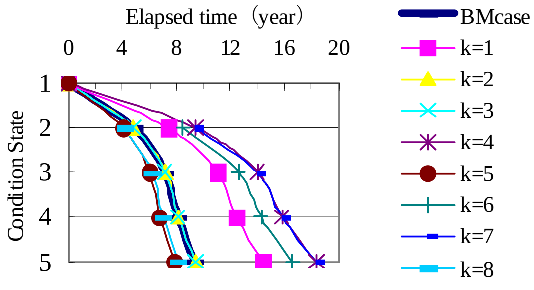

Figure 1: Expected deterioration path for each group

Table 3: Transition probability matrix for k=5

State	1	2	3	4	5
1	0.7805	0.1700	0.0328	0.0132	0.0035
2	0	0.5996	0.2136	0.1354	0.0514
3	0	0	0.2767	0.4334	0.2899
4	0	0	0	0.4062	0.5938
5	0	0	0	0	1

Result in "Fig. 1" enables us to have a good comparison of the dispersion of each group of highway with respect to the thickness of surface asphalt and the traffic volume around the average deterioration curve. The life expectancy of rating and deterioration curve are important indicators, which assist managers to select the desire representative highway for the life cycle cost analysis. In our study, the group with heteroneity factor ε = 1.1931 (k=5) is selected for sample application of the LCC analysis. Table V shows the results of transition probability for this group.

4. Conclusion

This paper has discussed the local mixture model for benchmarking the infrastructure technologies. The local mixing mechanism is expressed by mean of heterogeneity factor ε that exist in each group of infrastructure. Heterogeneity factor was estimated together with the markovian transition probability, which describes the deterioration process of the infrastructure. An empirical study has been conducted on the dataset of Vietnamese highway system, which was classified according to the technologies in asphalt pavement. The estimation results shown in Fig. 1 could be used as key performance indicators for selection of pavement design, technology offered from difference contractors.

References

1) Shin, H.C and Madanat, S.M, "Development of stochastic model of pavement distress initatation," Processing of JSCE Conference, vol. 744, pp. 61-67, 2003.

2) While, D.J, Markov Decision Process, Springer, 1992

3) Tsuda, Y. Kaito, K. Aoki, K and Kobayashi, K, Y. Tsuda, K. Kaito, K. Aoki and K. Kobayashi, "Estimating markovian transition probabilities for bridge deterioration forecasting," JSCE Journal (in Japanese), no. 801/I-73, October 2005.

4) Anaya, K. and Marriott, P, "Local mixtures of the exponential distribution," Annals of the Institute of Statistical Mathematics, vol. 59, no. 1, pp. 111-134, 2007.

5) TCN-22, Vietnamese Standards for Design of Highway- 22TCN-277-01, Ministry of Transportation, Vietnam, 2001.