HIDDEN MARKOV MODEL CONSIDERING THE INCONSISTENCIES OF OBSERVATION PERIOD

B. MIYAZAKI[1*†], K. OBAMA[2], and K. KAITO[3]

[1]*Department of Civil Engineering, Osaka University, Japan*
[2]*Department of Civil Engineering, Osaka University, Japan*
[3]*Department of Civil Engineering, Osaka University, Japan*

ABSTRACT

With the development of deterioration prediction of infrastructure, the model expressing a hierarchical relation between two different events have been proposed. However, the data acquired in the actual world, there are many cases where there are the inconsistencies of the observation period. Therefore, in previous studies, the estimation using all the data acquired was difficult. So the authors have estimated by intentionally truncating the data to compensate for inconsistency of the observation period. In this paper, the authors propose a hierarchical model that takes into account the inconsistency of data observation period. To express a complex deterioration process using a Poisson hidden Markov model is composed of localized damage such as potholes, and a damage of the entire pavement cracks. In the processes, the authors formulate the hierarchical model by using latent variables to compensate for inconsistency of the data observation period. So it is possible to perform the estimation using all the data acquired. Lastly, the validity of the proposed method is discussed through the empirical analysis of the visual inspection data.

Keywords: inconsistent data, hidden Markov deterioration model, pothole, crack, pavement

1. INTRODUCTION

In recent years, the development of deterioration prediction model of infrastructure is progressing rapidly (Lancaster, T. 1990; Mishalani, R. et al. 2002). In particular, for road pavement, the accumulation of data is advancing dramatically due to technological advances of the road inspection vehicle, so it has become possible to detailed analysis. At the same time, the model expressing a hierarchical relation between two different events have been proposed. For example, hierarchical hidden Markov deterioration model which represents the complex deterioration process that consists of lowering process of load bearing capacity and states of road surface (Kobayashi et al. 2011), and Poisson hidden Markov model which represents the deterioration process that consists of generating localized damages, which occur frequently, and the deterioration of the surface and base

* Corresponding author: Email:b-miyazaki@civil.eng.osaka-u.ac.jp

† Presenter: Email:b-miyazaki@civil.eng.osaka-u.ac.jp

layers of pavement, which progresses slowly (Nam et al. 2012). The quantitative evaluation of the relation between degradation event two different can lead to maintenance more reasonable is expected. However, the data acquired in the real world, there are many cases where there is inconsistency of the observation period. Therefore, in previous studies, the estimation is performed by using all the data acquired has been difficult. So the authors have estimated by intentionally truncating the data to compensate for inconsistency of the observation period. In order to take full advantage of the acquired data, building method considering inconsistencies observation period is required. In this study, the hierarchical model in consideration of the inconsistencies of the observation period of two different events is proposed. In that cases, the authors formulate the hierarchical model by using latent variables to compensate for inconsistencies in the data observation period, it is possible to perform the estimation using all data.

2. BASIC CONCEPTS OF THIS STUDY

2.1. Inconsistencies of observation period

For the description of the time inconsistency of data to be acquired, the authors takes up a method of checking two types of road patrol and road surface inspection as an example. The road administrator is conducting road surface inspection periodically. By road surface inspection, we're getting at the same time the three damage indicators; crack rate, rutting and IRI. Define the states of pavement using these indicators, the need for repair such a large overlay is considered. Further, it is not limited to the periodically road surface inspection, road administrator performs a road patrol routinely, thereby obtaining information about the generating of localized damage such as potholes. However, the inspection method changes with each data. Naturally, each inspection is independent and will not be carried out are synchronized. So, the data acquired in the real world, there are many cases where there is inconsistency of the observation period. So the authors have performed estimated by intentionally truncating the data to compensate for inconsistency of the observation period. The purpose of this study is to consider the inconsistency of the observation period. In previous studies, it was virtually set using latent variables, information about the state of the road surface in a localized point during the most recent road surface inspection and the next road surface inspection. In this study, the authors extend it. Concretely, in order to consider the effects of the potholes generated in the period outside of the most recent road surface inspection and the next road surface inspection, set the latent variable. The latent variables, each end point conditions are set. By letting the transition latent variables according to this condition, it is possible to obtain information about the state of road surface of not acquired normally.

2.2. Practical problems in the maintenance of highway pavement

In recent years, high performance road pavement has been introduced to the highway pavement. High performance road pavement is the drainage pavement which set up void ratio more highly. After starting introduction of high performance road pavement in 2003 and afterwards, the

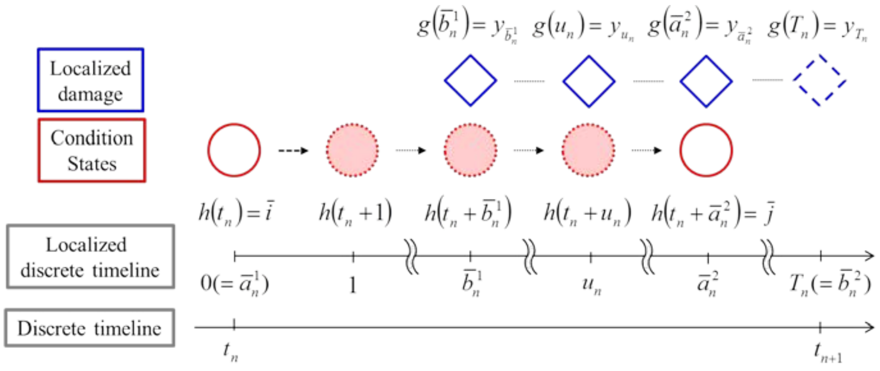

Figure 1: Localized discrete timeline.

generating of localized damage such as pothole increased, and it repaired more often for a short period of time repeatedly. A cause is depths damage to the basis and the subgrade by the osmosis function of rain water which high performance road pavement has. On the other hand, the state evaluation of the road in Japan, it is set as the control limits the step of crack rate acquired in the road surface inspection reached 20% in many cases. However, this is a reference value established when the main highway pavement was dense pavement. In the situations where frequent localized damage is represented by pothole, cases to carry out repairs or emergency measures in spite of not reached the reference value has increased. There are also reports crack rate before performing the repair in practice that it is around 10%. To fill the deviation from the practice, i.e. in order to carry out maintenance in line with actual conditions, it is necessary to express quantitatively the relation between the generating of localized damage acquired by daily road patrol and state evaluation by crack rate acquired in the road surface inspection. Poisson hidden Markov model which takes account of inconsistency time proposed in this study are to follow a Poisson hidden Markov model proposed by Nam et al expanded the theoretically. Moreover, there is a value in that it is in response to the request of practical to clarify the relation of the potholes and crack rate.

3. FORMULATION OF THE MODEL

As shown in Figure 1, the data about the road surface condition states is acquired in the period $[\bar{a}_n^1, \bar{a}_n^2](\bar{a}_n^1 = 0,1,\cdots,T_n-1; \bar{a}_n^2 = 1,2,\cdots,T_n; \bar{a}_n^1 < \bar{a}_n^2)$ localized discrete timeline u_n, and the data relating to localized damage have been acquired in the period $[\bar{b}_n^1, \bar{b}_n^2](\bar{b}_n^1 = 0,1,\cdots,T_n-1; \bar{b}_n^2 = 1,2,\cdots,T_n; \bar{b}_n^1 < \bar{b}_n^2)$. Information on road surface condition can be acquired at road surface inspections performed at points $t_n+\bar{a}_n^1$ and $t_n+\bar{a}_n^2$ $(n=1,2,\cdots)$. However, this information cannot be acquired at localized points $u_n(u_n = 0,\ldots, T_n)$. In other words, the condition at localized point u_n is a probability variable that cannot be observed by inspectors, but the authors hypothesize that state $h(t_n+u_n)=l_{u_n}^n$ is already known.

Now, let us examine the period $\tau_n = [t_n + t_{n+1})$, defined by the two inspection points, assuming we know the state $h(t_n+\bar{a}_n^1)=\bar{i}_n$, $h(t_n+\bar{a}_n^2)=\bar{j}_n$, for consecutive surface inspection points $t_n+\bar{a}_n^1, t_n+\bar{a}_n^2$. In this hypothesis, repair has not been performed on the pavement during period τ_n. We also know the number of potholes $g(u_n) = y_{u_n} \ (y_{u_n} = 0,\cdots,T_n)$ that were generated at localized points $u_n (u_n = \bar{b}_n^1,\cdots,\bar{b}_n^2)$, with inspection point $t_n + \bar{b}_n^1$ as the starting point. Although we do not have information regarding state $h(t_{n+}u_n)$ for localized points $u_n (u_n = \bar{b}_n^1,\cdots,\bar{b}_n^2)$, we know

$$
\begin{array}{ll}
h(t_{n-1} + T_{n-1} - 1) \le h(t_n) \le \cdots \le h(t_n + u_n) \le \cdots \le h(t_n + \bar{a}_n^1) = \bar{i}_n & (0 \le u_n \le \bar{a}_n^1) \\
\bar{i} = h(t_n + \bar{a}_n^1) \le \cdots \le h(t_n + u_n) \le \cdots \le h(t_n + \bar{a}_n^2) = \bar{j} & (\bar{a}_n^1 < u_n < \bar{a}_n^2) \\
\bar{j}_n = h(t_n + \bar{a}_n^2) \le \cdots \le h(t_n + u_n) \le \cdots \le h(t_n + T_n) \le h(t_{n+1} + 1) & (\bar{a}_n^2 \le u_n < T_n) \\
\begin{cases} h(t_{n-1} + T_{n-1} - 1) = 1 & when \quad n = 0 \\ h(t_{n+1} + 1) = I & when \quad n = N \end{cases} &
\end{array}
\qquad (1)
$$

holds true from the state $h(t_n + \bar{a}_n^1) = \bar{i}_n \ \ h(t_n + \bar{a}_n^2) = \bar{j}_n$. Let us now look at localized period $\iota_{u_n} = [u_n, u_{n+1})(u_n = 0,\cdots,T_n - 1)$, which composes period τ_n. The arrival rate of potholes $\mu(l_{u_n}^n, z_{u_n}) > 0$ in period ι_{u_n} corresponds to the Poisson generation model expressed by

$$
\mu(l_{u_n}^n, z_{u_n}) = z_{u_n} \alpha^{l_{u_n}^n} \qquad (2)
$$

However, $z_{u_n} = (z_{1,u_n},\cdots,z_{P,u_n})$ is the explanatory variable vector observed at localized point u_n. $\alpha^{l_{u_n}^n} = (\alpha_1^{l_{u_n}^n},\cdots,\alpha_P^{l_{u_n}^n})'$ is the unknown parameter vector, and also $\alpha = (\alpha^{1'},\cdots,\alpha^{I-1'})'$. P indicates the number of the explanatory variable. The arrival rate $\mu(l_{u_n}^n, z_{u_n})$ is defined according to localized period [u_n, u_n+1). Here, it is thought potholes arrive according to the Poisson generation model with average $\mu(l_{u_n}^n, z_{u_n})$ at point t. However, $l_{u_n}^n$ is the surface condition at localized point u_n, and is hypothesized to be a certain value within localized period ι_{u_n}. Because the localized period ι_{u_n} is standardized as 1, the conditional probability $\pi(y_{u_n} \mid l_{u_n}^n, z_{u_n})$ that a y_{u_n} number of potholes are generated within localized period ι_{u_n}, can be expressed as:

$$
\begin{aligned}
\pi\left(y_{u_n} \mid l_{u_n}^n, z_{u_n}\right) &= \mathrm{Prob}[\mathrm{g(u_n)} = y_{u_n} \mid h(t_n + u_n) = l_{u_n}^n, z_{u_n}] \\
&= \exp\{-\mu(l_{u_n}^n, z_{u_n})\} \frac{\{\mu(l_{u_n}^n, z_{u_n})\}^{y_{u_n}}}{y_{u_n}!}
\end{aligned}
\qquad (3)
$$

The explanation about the Markov deterioration hazard model which estimates transition between states is yielded to reference (Tsuda et al. 2006).

Table 1: Data specifications

Month which conducted road surface inspection	2009/7	
The number of section acquired crack rate	1,671	
Repair history data acquisition period	1992/6~2009/7	
Road patrol data acquisition period	2007/8~2011/9	
The number of section generating potholes	128	
The total number of generating potholes	332	
The number of generating potholes by period	2007/8~2009/6	92
	2009/7~2011/9	240

Table 2: Definition of States

States	Crack rate Cr(%)	Sample
1	Cr=0	645
2	0<Cr<2.5	731
3	2.5≦Cr<5	161
4	5≦Cr<10	101
5	10≦Cr<20	29
6	20≦Cr	4

4. CASE STUDY

4.1. The outline of an analysis data

The Poisson hidden Markov model proposed in this paper was applied to the maintenance of road pavement, targeting part of expressway, in order to empirically investigate its validity. The targeted roads have data on past surface inspections and repairs, as well as daily road patrols. Each data specifications are shown in a table 1. Timeline data on the pavement deterioration is necessary for estimating the Markov deterioration hazard model. For this case study, surface inspection data of the year 2009, and the repair data from June 1992 to July 2009 was used. Moreover, the acquired data was arranged every 100m. In the multi-staged exponential Markov model, it was assumed that the roads' damages were eliminated completed during repairs, therefore the change and transition of states from the most recent repair to the inspection points of 2009 were used. States of pavement was defined using the crack rate of pavement acquired during surface inspections. Table 2 shows the definition of state evaluation using crack rate. States are evaluated in a scale of six, from 1 to 6. A State of over Cr 20%, state in which deterioration advanced most, was given as 6.

During daily patrols, as a basic rule one road section was patrolled each day, surface abnormalities and obstacles were inspected visually, and when something was observed, the content, time and position was recorded. In this case, the information on pothole to be used in the study was acquired from the daily patrols, so the Poisson generation model will be estimated using this information. Moreover, the authors used information on road patrols of expressway from June 2007 to September 2011. For these road sections, the authors created basic road sections of 100m each. As a result, the authors had a total of 1,671 basic sections that could be used for Poisson hidden Markov

Table 3: Estimation result (Parameters of the multi staged exponential Markov model)

Condition States	Constant term $\beta_{i,1}$	Heavy Vehicle Traffic Volume $\beta_{i,2}$	Average Hazard rate $E[\lambda(i)]$	Expected Life Span $E[RMD]$(year)
1	-3.896 (-3.958, -3.834) 0.003	0 (-) -	0.02	4.100
2	-5.556 (-5.958, -5.169) 0.09	0.912 (0.306, 1.53) -0.093	0.007	12.334
3	-4.751 (-5.467, -4.186) -0.109	1.166 (0.322, 2.242) 0.116	0.018	4.721
4	-6.446 (-7.241, -5.589) -0.103	2.714 (1.535, 3.813) 0.116	0.008	9.968
5	-5.397 (-6.142, -4.663) 0.068	0 (-) -	0.005	18.388

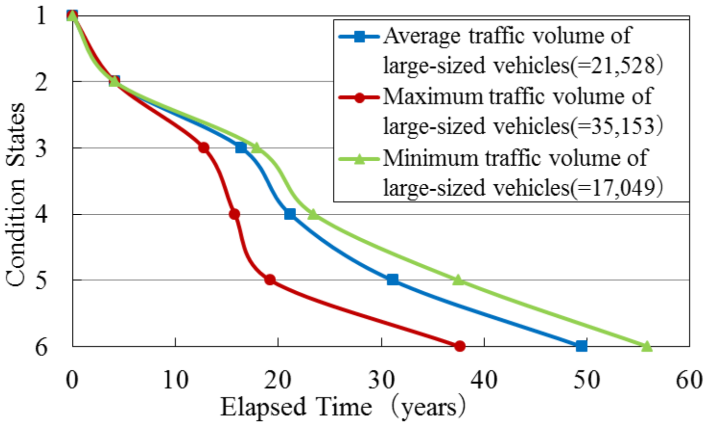

Figure 2: Expected deterioration path of road surface.

model estimation. Also, 128 of these basic sections had been confirmed to include potholes. When one or more occurrences in the one section included, the total generating number of potholes became 332 pieces. It noted that the observed period of states of road surface, observation period of potholes in the form of a road patrol is not consistent. In previous studies, it is impossible to be used for estimation for the data of the potholes, which is acquired after July 2009, it will be limited to the data through from June 2007 to July 2009. Total number of generating potholes in the period was 92 pieces, so data about the potholes of 234 pieces were not available to estimate. In this case, it is conceivable that would underestimate the probability of the potholes. The authors propose a method that allows estimation using all of the data acquired in this study.

4.2. Bayesian estimation of the multi-staged exponential Markov model

Table 3 shows the results of Bayesian estimation on the multi-staged exponential Markov model using state defined by crack rate (parameter expectation values and 90% credible interval (Bayesian

Table 4: Estimation result (Parameters of the Poisson model)

Condition States	Constant term $\alpha_{i,1}$	Traffic Lane $\alpha_{i,2}$	Road Structure $\alpha_{i,3}$	Average arrival rate $E[\mu(i)]$
1	-9.676 (-10.488, -8.889) -0.144	5.046 (4.252, 5.857) 0.173	1.041 (0.719, 1.352) -0.037	0.028
2	-7.63 (-7.97, -7.326) -0.035	0 (-) -	2.052 (1.604, 2.518) 0.102	0.004
3	-8.653 (-9.373, -8.037) -0.138	4.045 (3.404, 4.749) 0.113	0 (-) -	0.010
4	-11.8 (-12.768, -10.761) -0.055	3.474 (2.605, 4.265) -0.056	6.373 (5.481, 7.518) 0.098	0.142
5	-9.682 (-10.854, -8.527) 0.013	0 (-) -	3.879 (2.233, 5.475) 0.08	0.003
6	-3.833 (-4.571, -3.253) -0.009	0 (-) -	0 (-) -	0.022

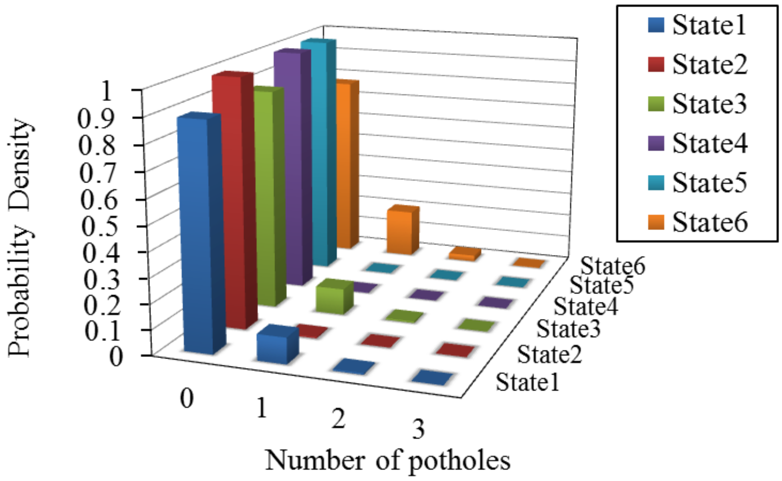

Figure 3: Poisson distribution of each state (After 1 year).

confident interval)). Furthermore, Geweke statistics are also shown. These characteristic variables were decided based on coinciding conditions and Geweke statistics (Geweke, 1996). As a result, only heavy vehicle traffic volume was selected to be used in this analysis. The heavy vehicle traffic volume in the target road sections is an average 24,184/day, minimum 17,049/day, and maximum 35,153/day. For estimation, these values were standardized so the maximum was 1. Table 3 also highlights the values of expected hazard rate $E[\lambda_i]$ and the expected lifetime $E[RMD]=1/\lambda_i$ of each state i. From this we can see that the expected lifetime of states 1 and 3 are relatively short. Figure 2 shows the expected deterioration path of road surface was calculated. Moreover, the path includes variance due to the heavy vehicle traffic volume. Road sections with the most traffic have expected pavement lifetime of approximately 37.6 years, and those with the least traffic at approximately 55.9 years. Because repairs are often conducted at state 5, as defined in this study, the expected lifetime is an average of approximately 31.1 years, maximum approximately 19.2 years and minimum approximately 37.5 years.

4.3. Bayesian estimation of Poisson generation model

The results of Bayesian estimation on the Poisson generation model are shown in Table 4. In the Poisson generation function of the Poisson hidden Markov model, the characteristic variables defined in arrival rate function are traffic lane(slow lane or passing lane) and road structure(bridge or embankment), respectively. As can be seen from the table, mixture Poisson generation models can be created for each state (in this case study, six Poisson generation models can be defined).

Figure.3 shows Poisson distribution of each state, at the point of 1 year after emergency repairs using potholes patching mixtures. Characteristic variable is slow lane and embankment. From the figure, in the states of 1, 3 and 6, the probability that pothole occurs one is high. High performance road pavement becomes to be used for the main paved road, the occurrence of pothole without crack increased. Therefore, at low crack rate such as state1 and state3, result the probability of occurrence of pothole is high has appeared. Further, in the state 6(more than 20% crack rate), deterioration has progressed significantly so it led to increase the probability of occurrence of pothole.

5. CONCLUSION

The authors have proposed a hierarchical model that allows the estimation using all the data to take into account the inconsistencies of the observation period of the acquired data. By presenting quantitatively the relation between the two events deterioration, can lead to maintenance more reasonable is expected. In this study, it covers the road surface inspection data, and data of daily road patrols, it is necessary to expand the coverage further.

REFERENCES

Kobayashi, K., Kaito, K., Eguchi, T., Ohi, A., and Okizuka, R. (2011).A hierarchical hidden Markov deterioration model for pavement structure, Infrastructure Planning and Management, JSCE, Vol.67, No.4, 422-440.(in Japanese)

Nam, L. T., Kaito, K., Kobayashi, K., and Okizuka, R. (2012). Modeling pavement deterioration processes by Poisson hidden Markov models, Infrastructure Planning and Management, JSCE, Vol.68, No.2, 62-79.(in Japanese)

Lancaster, T. (1990). The Econometric Analysis of Transition Data, Cambridge University Press.

Mishalani, R. and Madanat, S. (2002). Computation of infrastructure transition probabilities using stochastic duration models, ASCE Journal of Infrastructure Systems, Vol.8, No.4.

Tsuda, Y., Kaito, K., Aoki, K., and Kobayashi, K. (2006).: "Estimating markovian transition probabilities for bridge deterioration forecasting." Journal of Structural Engineering and Earthquake Engineering, JSCE, 23(2), 241–256.

Geweke, J. (1996). "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments." Research Department Staff Report 148, Federal Reserve Bank of Minneapolis.