International Journal of Architecture, Engineering and Construction
Vol.1, No. 1, March 2012, pp. 1-13

# A Bayesian Estimation Method to Improve Deterioration Prediction for Infrastructure System with Markov Chain Model

Kiyoshi Kobayashi[1,*], Kiyoyuki Kaito[2], Nam Lethanh[3]

[1]Graduate School of Management, Kyoto University, Kyoto 606-8501, Japan

[2]Department of Civil Engineering, Osaka University, Suita, Osaka 565-0871, Japan

[3]Institute of Construction and Infrastructure Management, Swiss Federal Institute of Technology (ETH), Zurich 8093, Switzerland

**Abstract:** In many practices of bridge asset management, life cycle costs are estimated by statistical deterioration prediction models based upon monitoring data collected through inspection activities. In many applications, it is, however, often the case that the validity of statistical deterioration prediction models is flawed by an inadequate stock of inspection dates. In this paper, a systematic methodology is presented to provide estimates of the deterioration process for bridge managers based upon empirical judgments at early stages by experts, and whereby revisions may be made as new data are obtained through later inspections. More concretely, Bayesian estimation methodology is developed to improve the estimation of Markov transition probability of the multi-stage exponential Markov model by Markov chain Monte Carlo method using Gibbs sampling. The paper concludes with an empirical example, using the real world monitoring data, to demonstrate the applicability of the model and its Bayesian estimation method in the case of incomplete monitoring data.

**Keywords:** Bayesian estimation, Markov chain Monte Carlo, Gibbs sampling, statistical deterioration prediction, Markov chain model, infrastructure management

**DOI:** 10.7492/IJAEC.2012.001

## 1 INTRODUCTION

In the asset management for civil infrastructures, it is important to determine optimal intervention strategies based on the life cycle costs analysis (LCCA) (Kobayashi and Kuhn 2007; Kobayashi and Ueda 2003). Life cycle costs are the sum of all expenses incurred to stakeholders (owners, users, and the public) during the lifetime of civil infrastructures. The LCCA is often carried out either with minimization of discount flow method (Jido et al. 2008) or minimization of annual average cost method (Kobayashi and Kuhn 2007; Kaito et al. 2005). One important, but subtle, feature of the LCCA is that it has to be dependent on the deterioration model, which is employed to predict the deterioration of civil infrastructure. The deterioration model is therefore considered as the heart of an infrastructure management system (IMS).

To the authors's knowledge, so far in infrastructure asset management study area, researchers employs two kinds of deterioration prediction methods: 1) one targets the average deterioration of all related civil infrastructures, and 2) the other targets the concrete damage and deterioration of each infrastructure or member. For the former, a statistical method is often adopted in order to model the regularities behind deterioration processes based on an enormous amount of deterioration information. For the latter, a physical method is frequently adopted in order to directly model deterioration processes by clarifying the deterioration mechanism. In addition, when surveying the actual condition state of infrastructure asset, it is common at the outset to grasp the deterioration characteristics of the entire civil infrastructures, make decisions on intervention strategies (such as the selection of methods for preventive maintenance or posteriori maintenance),

*Corresponding author. Email: kobayashi.kiyoshi.6n@kyoto-u.ac.jp

and then design concrete intervention plans for each infrastructure. In favor of statistical method in deterioration prediction, this study aims to improve the application of statistical deterioration prediction models by introducing a novel Bayesian estimation method.

We have developed several statistical models and the estimation methods (Kaito et al. 2002; Tsuda et al. 2006b; Kobayashi et al. 2010) based on monitoring data, conducted many empirical studies using actual data, and verified their effectiveness. One important finding from our previous work is that, in order to assure a high accuracy in deterioration prediction, it is necessary to collect several thousands of monitoring data (Sugisaki et al. 2006; Tsuda et al. 2006b). However, in practice, most infrastructure administrators face the problems of insufficient amounts of data, and thus, it is anticipated that the insufficiency of data will hinder the practical application of the statistical method.

Suffice it to say that in order to improve the applicability of statistical deterioration forecasting models, it is essential to develop a methodology which can provide estimation results by fusing prior information, such as the experience, knowledge, and know-how of professional engineers, with a limited amount of monitoring data, accumulate additional data, and update the estimation results constantly. Under this circumstance, this study aims to develop an estimation method that updates parameters of deterioration prediction model constantly based on Bayesian statistical approach. Precisely, the method is developed by employing Bayesian estimation and Markov chain Monte Carlo (MCMC) for numerically estimating Markov transition probability (mtp) of the multi-stage exponential Markov (MUSTEM) model, which was developed by Tsuda et al. (2006b). Our favor for improving estimation method for the MUSTEM model is mainly due to the reason that the MUSTEM model has advantageous features that the model is capable of handling the heterogeneity of data of individual infrastructure, including structural characteristics, environmental conditions, inspection intervals, in a non-aggregative manner. Moreover, it allows transition probabilities with more than 2 steps, while it is not possible with other developed models used in the field.

Following section gives a review on research background and clarifies the contribution of this study. Section 3 describes briefly the formulation of the MUSTEM model and its numerical estimation approach using maximum likelihood estimation (MLE) approach. In section 4, our methodology using Bayesian estimation and MCMC method is elaborated in detail. Section 5 discusses an empirical study using monitoring data of reinforced concrete slabs (hereinafter abbreviated as "RC slabs") of bridges. The last section summaries the paper with recommendations for future research works.

## 2 RESEARCH BACKGROUND

Statistical methods differ from physical methods in that a statistical method models the regularities existing behind deterioration processes without clarifying the deterioration mechanism. Therefore, statistical methods are useful for discussing the average deterioration characteristics at a macroscopic level of management. However, in order to secure precision in prediction, it is essential to accumulate as many as several thousands of monitoring data. In order to overcome this problem, this study provides prediction results by fusing the prior information of professional engineers with a limited amount of monitoring data at the early stage when monitoring data has not been accumulated to a sufficient degree. Furthermore, this study develops a methodology based on Bayesian estimation, which can constantly update prediction results as data is accumulated. Especially, this study presents a methodology, while focusing on a disaggregative method, which is considered as one of the most advanced statistical methods to date. The proposed method is able to secure high precision in prediction using a small number of data samples compared with those of other methods.

In Bayesian statistics, the occurrence probability of an event is estimated subjectively by utilizing the information, knowledge, experiences, etc. possessed by professional engineers. In addition, prior information can be used in a proactive way, and thus it is possible to estimate model's parameters even when samples are few. Moreover, Bayesian estimation has an excellent feature that it can update the model's parameters readily whenever new data samples arrive (Ibrahim et al. 2001; Jeff 2006).

However, as mentioned later in section 4, when Bayesian estimation is applied to estimate the parameters of the MUSTEM model, there is no conjugate relation between the prior distribution and the posterior distribution about the model's parameters (The two distributions are different in category). Moreover, the likelihood function is in the form of multiple-integration that makes the numerical analysis a great difficulty when applying the MLE method. Therefore, it is impossible to obtain the posterior distribution of model's parameters analytically. This issue has been considered as an obstacle to practical application of the MLE method (Jeff 2006).

In the past few years, the MCMC method has been developed and significantly advanced. The method has a great feature that it can be used to randomly generate samples of model's parameters so that the mean values of model's parameters could reach to a convergent space (Gamerman and Lopes 2006). An example of Bayesian estimation application with the MCMC method can be referred to Tsuda et al. (2006a). In the cited paper, the author developed a method for estimating parameters of a Weibull deterioration hazard model targeted at the facilities and devices whose deterioration status can be described, based on the Bayesian estimation, by whether or not they have any failure. In this study, the authors expand this method and propose a methodology for conducting the Bayesian estimation of the mtp based on the MUSTEM model.

## 3 MULTI-STAGE EXPONENTIAL MARKOV (MUSTEM) MODEL

In this section, we summary the research of Tsuda et al. (2006b), which proposed a modeling approach to estimate the mtp based on monitoring data. Careful readers are recommended to refer to the original paper for greater details of the methodology.

### 3.1 The Model

It is assumed that the condition states of a civil structure is provided as monitoring data and its history performance path can be shown as in Figure 1. In the figure, time $\tau$ represents the actual time on a calendar (hereafter refered as "time"). Deterioration of the structure can be represented by discrete condition state $i(i = 1, .., J)$, with $i = 1$ as initial condition state (when structure is new) and $i = J$ as absorbing condition state. Time $\tau_A$ and $\tau_B$ are actual inspection times, while time $\tau_i$ is any arbitrary time in between. Duration between two inspection times is $Z$. Given monitoring data of two inspection times $\tau_A$ and $\tau_B$, the mtp can be described as follows:

$$\text{Prob}[h(\tau_B) = j \mid h(\tau_A) = i] = p_{ij} \quad (1)$$

Suppose that the condition state changes from $i$ to $i+1$ at time $\tau_i$ (timing $y_C$). At that time, the duration of being in condition state $i$ can be expressed by the following equation: $\zeta_i = \tau_i - \tau_{i-1} = y_C$. The duration $\zeta_i$ of condition state $i$ is considered as a random variable, and is subjected to the probability density function $f_i(\zeta_i)$ and the distribution function $F_i(\zeta_i)$. Here, the domain of the duration $\zeta_i$ is $[0, \infty)$. The following expression can be derived from the definition of distribution function:

$$F_i(y_i) = \int_0^{y_i} f_i(\zeta_i) d\zeta_i \quad (2)$$

The distribution function $F_i(y_i)$ represents the cumulative probability of the change of the condition state from $i$ to $i+1$ in period from initial timing $y_i = 0$ (time $\tau_{i-1}$), at which the condition state has become $i$, to the timing $y_i$ (time $\tau_{i-1} + y_i$). Accordingly, the probability $\tilde{F}_i(y_i)$ of remaining at condition state $i$ from the initial timing $y_i = 0$ to the sample timing $y_i \in [0, \infty)$ can be expressed by the following equation, using the cumulative probability of the change of the condition state from $i$ to $i+1$ until timing $y_i$:

$$\text{Prob}\{\zeta_i \geq y_i\} = \tilde{F}_i(y_i) = 1 - F_i(y_i) \quad (3)$$

The conditional probability of the event that the structure remains in condition state $i$ until timing $y_i$ and change to condition state $i+1$ in the period $[y_i, y_i + \Delta y_i)$ can be defined as:

$$\lambda_i(y_i)\Delta y_i = \frac{f_i(y_i)\Delta y_i}{\tilde{F}_i(y_i)} \quad (4)$$

The instantaneous rate $\lambda_i(y_i)$ of the change in the condition state of the target structure from $i$ to $i+1$ at timing $y_i$ is called a hazard function. By using a hazard function suited for the assumed deterioration process, it is possible to describe initial damage, accidental deterioration, and deterioration with time, etc.

Under the assumption that the Markov characteristics concerning the deterioration processes of civil in-

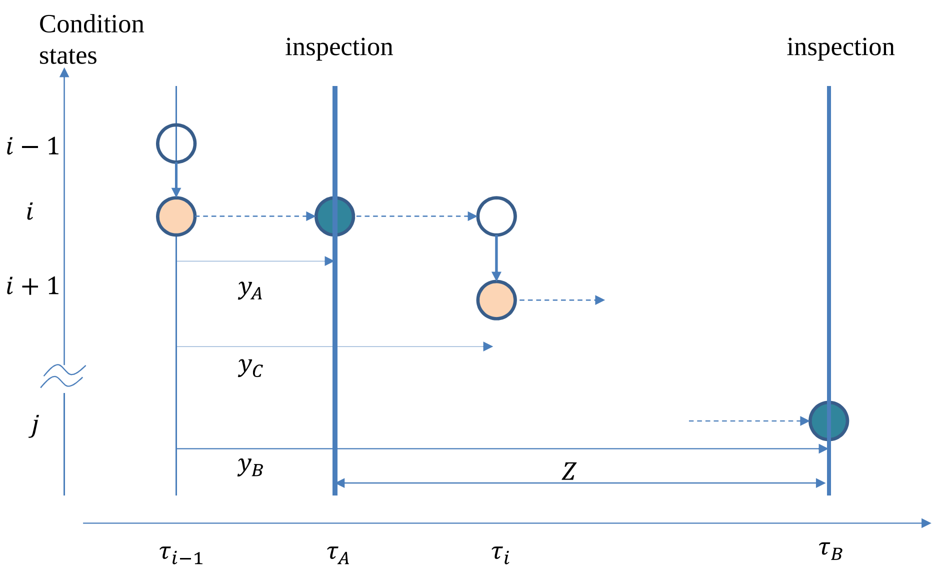



**Figure 1.** Deterioration process and inspection time

frastructures do not depend on the history of deterioration and the hazard function is constant $\theta_i > 0$, in another words, the hazard function is independent of the timing $y_i$, following equation is defined:

$$\lambda_i(y_i) = \theta_i \tag{5}$$

Using the hazard function $\lambda_i(y_i) = \theta_i$, the probability of the event that condition state $i$ remains over a duration $y_i$ can be further described as the survival probability function:

$$\tilde{F}_i(y_i) = \exp\left[-\int_0^{y_i} \lambda_i(u)du\right] = \exp(-\theta_i y_i) \tag{6}$$

The survival probability function is identical to the mtp $p_{ii}$ when the duration $y_i$ equals to interval $z$. By defining the subsequent conditional probability of condition state $j$ to $i$, with respect to $z$, a general mathematical formula for estimating the mtp $p_{ij}$ is defined as:

$$\begin{aligned} p_{ij}(z) =& \text{Prob}[h(\tau_B) = j | h(\tau_A) = i] \\ =& \sum_{k=i}^{j} \prod_{m=i}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} \cdot \\ & \prod_{m=k}^{j-1} \frac{\theta_m}{\theta_{m+1} - \theta_k} \exp(-\theta_k z) \end{aligned} \tag{7}$$

where there are the following conditions:

$$\begin{cases} \displaystyle\prod_{m=i}^{k-1} \frac{\theta_m}{\theta_m - \theta_k} = 1 & \text{at} \quad (k \leq i+1) \\ \displaystyle\prod_{m=k}^{j-1} \frac{\theta_m}{\theta_{m+1} - \theta_k} = 1 & \text{at} \quad (k \geq j) \end{cases} \tag{8}$$

In addition, the mtp $p_{iJ}$ for any condition state $i$ to go to the absorbing condition state $J$ is defined as:

$$p_{iJ}(z) = 1 - \sum_{j=i}^{J-1} p_{ij}(z) \ (i = 1, \cdots, J-1) \tag{9}$$

## 3.2 Estimation Approach Using the MLE Method

The MLE method is a popular statistical estimation method to obtain the most likely value of model's parameter based on monitoring data. To use the MLE method, hazard rate $\theta_i$ is a function of characteristic variables $\boldsymbol{x}^k$ and unknown parameters $\boldsymbol{\beta}_i = (\beta_{i,1}, \cdots, \beta_{i,M})$ where $M$ is number of characteristic variables and $k(k = 1, ..., K)$ is index of monitoring data.

$$\theta_i^k = f(\boldsymbol{x}^k : \boldsymbol{\beta}_i) \tag{10}$$

Under the assumption of hazard function in Eq. (10), the mtp can be expressed by $p_{ij}(\bar{z}^k, \bar{\boldsymbol{x}}^k : \boldsymbol{\beta})$ as a function of actual measurement data $(\bar{z}^k, \bar{\boldsymbol{x}}^k)$ and unknown parameters $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_{J-1})$. Assuming that the deterioration phenomena of $K$ civil infrastructures are independent of one another, the likelihood function representing the probability density of the simultaneous occurrence of the deterioration transition pattern of all inspection samples can be formulated as follows (Tobin 1958; Amemiya and Boskin 1974):

$$\mathcal{L}(\boldsymbol{\beta}) = \prod_{i=1}^{J-1} \prod_{j=i}^{J} \prod_{k=1}^{K} \left\{ p_{ij}(\bar{z}^k, \bar{\boldsymbol{x}}^k : \boldsymbol{\beta}) \right\}^{\bar{\delta}_{ij}^k} \tag{11}$$

where $\delta_{ij}^k$ is dummy variable, having its value of 1 when $h(\tau_A^k) = i$ and $h(\tau_B^k) = j$ and 0 otherwise. Employing the MLE approach for the log likelihood function of Eq. (11), parameter value $\boldsymbol{\beta}_i$ can be estimated. As a result, hazard rate $\theta_i$ and the mtp $p_{ij}$ can be obtained.

The expected duration of condition state $i$ can be defined by means of survival function $\tilde{F}_i(y_i^k)$ (Lancaster 1990):

$$\begin{aligned} R_i^k &= \int_0^\infty \tilde{F}_i(y_i^k) dy_i^k \\ &= \int_0^\infty \exp(-\theta_i^k y_i^k) dy_i^k = \frac{1}{\theta_i^k} \end{aligned} \tag{12}$$

# 4 BAYESIAN ESTIMATION METHOD FOR THE MUSTEM MODEL

## 4.1 Bayes's Theorem

Although literature on Bayesian statistics has been documented in a countless number, we would prefer to briefly present the Bayes's theorem in this section for the convenience of the readers and to serve as connection to the later subsections of this section.

In Bayesian statistics, posterior distribution of parameters is estimated by using a likelihood function given observed data and an assumed prior distribution of model's parameters. Here, the likelihood function is represented by $\mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\xi})$. $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$ denote unknown parameter vector and observed data, respectively. It is assumed that $\boldsymbol{\beta}$ is a random variable, and is subjected to prior probability density function $\pi(\boldsymbol{\beta})$. Under these conditions and according to the Baye's law, when the observed data $\boldsymbol{\xi}$ is given, the posterior probability density function $\pi(\boldsymbol{\beta}|\boldsymbol{\xi})$ of unknown parameters $\boldsymbol{\beta}$ is defined as:

$$\pi(\boldsymbol{\beta}|\boldsymbol{\xi}) = \frac{\mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\xi})\pi(\boldsymbol{\beta})}{\int_\Theta \mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\xi})\pi(\boldsymbol{\beta})d\boldsymbol{\beta}} \tag{13}$$

where $\Theta$ represents the parameter space. At this time, $\pi(\boldsymbol{\beta}|\boldsymbol{\xi})$ can be expressed as follows:

$$\pi(\boldsymbol{\beta}|\boldsymbol{\xi}) \propto \mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\xi})\pi(\boldsymbol{\beta}) \tag{14}$$

The symbol $\propto$ denotes "be proportional to". The denominator of the right-hand side of Eq. (13):

$$m(\boldsymbol{\xi}) = \int_\Theta \mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\xi})\pi(\boldsymbol{\beta})d\boldsymbol{\beta} \tag{15}$$

is called the normalization constant of $\pi(\boldsymbol{\beta}|\boldsymbol{\xi})$, or the prior predictive distribution. In general, the procedures of the Bayesian estimation can be summarized in following steps:

Step 1: the prior probability distribution function $\pi(\boldsymbol{\beta})$ is specified, based on the prior information;
Step 2: the likelihood function $\mathcal{L}(\boldsymbol{\beta}|\boldsymbol{\xi})$ is defined based on the newly obtained data $\boldsymbol{\xi}$;
Step 3: the prior probability density function is modified in accordance with the Bayes' theorem, and the posterior probability density function $\pi(\boldsymbol{\beta}|\boldsymbol{\xi})$ regarding the parameters $\boldsymbol{\beta}$ is updated.

The above steps are regarded as "Bayesian estimation rule" in our study. Theoretically, the Bayesian estimation method differs from the MLE method in that the probability distribution of the unknown parameters $\boldsymbol{\beta}$ is obtained as a posterior distribution. Moreover, in numerical solution with the MLE approach, to obtain the most likely value of model's parameters, it is required to derive the Jacobian (first derivative) and Hessian (second derivative) matrices for the objective function (e.g. Eq. (11)). Meanwhile, with Bayesian estimation method, Jacobian and Hessian matrices are not of requirement, and this feature is also considered as one of the advantage of Bayesian estimation method over the MLE method.

## 4.2 Bayesian Estimation Rule

This section presents our methodology used to estimate the parameter vector $\boldsymbol{\beta}$ of the MUSTEM model with the Bayesian estimation rule, using monitoring data. The newly obtained data is represented by $\bar{\boldsymbol{\xi}} = (\bar{\boldsymbol{\xi}}^1, \cdots, \bar{\boldsymbol{\xi}}^K)$. Also, it is assumed in our methodology that the available information for the inspection sample $k$ is $\bar{\boldsymbol{\xi}}^k = (\bar{\delta}^k_{ij}, \bar{z}^k, \bar{\boldsymbol{x}}^k)$.

By substituting the mtp $\pi_{ij}(z)$ into Eq. (11), the likelihood function becomes:

$$\mathcal{L}(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}) = \prod_{i=1}^{J-1}\prod_{j=i}^{J}\prod_{k=1}^{K}\Big\{\sum_{h=i}^{j}\prod_{l=i}^{h-1}\frac{\theta^k_l}{\theta^k_l-\theta^k_h} \prod_{l=h}^{j-1}\frac{\theta^k_l}{\theta^k_{l+1}-\theta^k_h}\exp(-\theta^k_h\bar{z}^k)\Big\}^{\bar{\delta}^k_{ij}} \tag{16}$$

where hazard rate $\theta_i$ is defined as a function of characteristic vector $\boldsymbol{x}^k$ and unknown parameter vector $\boldsymbol{\beta}_i$. As previously mentioned, the first important step in Bayesian estimation method is to define the prior probability density function for unknown parameter vector $\boldsymbol{\beta}$. However, it is not easy to select arbitrary probability distribution to represent the prior distribution of the parameter as it might result in computational problem or no convergence of estimation results (Ibrahim et al. 2001). To overcome this problem, it is important to define a conjugate prior probability density function for the parameter so that the posterior probability density function can be derived in similar form, which eases the convergence of the parameter in numerical computation.

In our methodology, we assume the prior probability density function of the parameter $\boldsymbol{\beta}_i$ follows the conjugate multidimensional normal distribution $\boldsymbol{\beta}_i \sim \mathcal{N}_M(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, which satisfies to derive a similar function form for the posterior probability density function as well. Given that assumption, we further express the probability density function of the $M$-dimensional normal distribution $\mathcal{N}_M(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$as:

$$g(\boldsymbol{\beta}_i|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{1}{(2\pi)^{\frac{M}{2}}\sqrt{|\boldsymbol{\Sigma}_i|}}\cdot \exp\left\{-\frac{1}{2}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_i)\boldsymbol{\Sigma_i}^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_i)'\right\} \tag{17}$$

where $\boldsymbol{\mu}_i$ represents the prior expectation vector (or mean) of $\mathcal{N}_M(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, and $\boldsymbol{\Sigma}_i$ is the prior variance-covariance matrix. The symbol $'$ denotes transposition. Based on Eqs. (14) and (16), the posterior probability density function $\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}})$ is defined as:

$$\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}) \propto \mathcal{L}(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}})\prod_{i=1}^{J-1} g(\boldsymbol{\beta}_i|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \propto \prod_{i=1}^{J-1}\prod_{j=i}^{J}\prod_{k=1}^{K}\Big\{\prod_{l=i}^{j-1}\theta^k_l\sum_{h=i}^{j}\prod_{l=i}^{h-1}\frac{1}{\theta^k_l-\theta^k_h}\cdot \prod_{l=h}^{j-1}\frac{1}{\theta^k_{l+1}-\theta^k_h}\exp(-\theta^k_h\bar{z}^k)\Big\}^{\bar{\delta}^k_{ij}}\cdot \prod_{i=1}^{J-1}\exp\Big\{-\frac{1}{2}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_i)\boldsymbol{\Sigma_i}^{-1}(\boldsymbol{\beta}_i-\boldsymbol{\mu}_i)'\Big\} \tag{18}$$

However, it is extremely hard to define the following normalization constant analytically, and therefore it is highly recommended to obtain the multiple integration value through numerical calculation.

$$m(\bar{\boldsymbol{\xi}}) = \int_{\Phi}\mathcal{L}(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}})\prod_{i=1}^{J-1} g(\boldsymbol{\beta}_i|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)d\boldsymbol{\beta} \tag{19}$$

Generally, a great deal of statistic literature has concluded that it is hard to obtain the posterior probability density function $\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}})$ of the parameter vector $\boldsymbol{\beta}$ if using the MLE method. Following sections present in greater detail of our method to overcome such limitation by using Gibbs sampling algorithm in the MCMC simulation to directly obtain the statistical values regarding the posterior distribution of parameters.

## 4.3 Gibbs Sampling

This section describes the Gibbs sampling, which is an algorithm in the MCMC method (Gamerman and Lopes 2006). The Gibbs sampling method is a method for obtaining samples from the posterior distribution by generating randomly samples of parameters $\boldsymbol{\beta}$ repeatedly, using the conditional posterior probability density function. Particularly, when it is difficult to directly obtain the posterior probability density function $\pi(\boldsymbol{\beta}|\boldsymbol{\xi})$. In order to explain the algorithm of the Gibbs sampling, observed data and unknown parameter vector are represented by $\bar{\boldsymbol{\xi}}$ and $\boldsymbol{\beta}$. In addition, unknown parameter vector subtracting $\beta_{e,m}$ from $\boldsymbol{\beta}$ is denoted by $\boldsymbol{\beta}^{-(e,m)}$. At this time, based upon the Eq. (18), the conditional posterior probability density function $\pi(\beta_{e,m}|\boldsymbol{\beta}^{-(e,m)}, \bar{\boldsymbol{\xi}})$ of $\beta_{e,m}$ when $\boldsymbol{\beta}^{-(e,m)}$ is known as follows:

$$\pi(\beta_{e,m}|\boldsymbol{\beta}^{-(e,m)}, \bar{\boldsymbol{\xi}}) \propto \prod_{i=1}^{e}\prod_{j=e}^{J}\prod_{k=1}^{K}\Big\{\theta_e^{k\,\bar{\delta}_{ij}^k-\bar{\delta}_{ie}^k}\cdot$$

$$\sum_{h=i}^{j}\prod_{l=i}^{h-1}\frac{1}{\theta_l^k-\theta_h^k}\prod_{l=h}^{j-1}\frac{1}{\theta_{l+1}^k-\theta_h^k}\exp(-\theta_h^k\bar{z}^k)\Big\}^{\bar{\delta}_{ij}^k}\cdot \tag{20-a}$$

$$\exp\Big\{-\frac{1}{2}(\boldsymbol{\beta}_e-\boldsymbol{\mu}_e)\boldsymbol{\Sigma_e}^{-1}(\boldsymbol{\beta}_e-\boldsymbol{\mu}_e)'\Big\}$$

$$\propto\prod_{i=1}^{e}\prod_{j=e}^{J}\prod_{k=1}^{K}\Big\{\exp\left(\beta_{e,m}x_m^k\right)^{\bar{\delta}_{ij}^k-\bar{\delta}_{ie}^k}\cdot$$

$$\sum_{h=i}^{j}\prod_{l=i}^{h-1}\frac{1}{\theta_l^k-\theta_h^k}\prod_{l=h}^{j-1}\frac{1}{\theta_{l+1}^k-\theta_h^k}\exp(-\theta_h^k\bar{z}^k)\Big\}^{\bar{\delta}_{ij}^k}\cdot \tag{20-b}$$

$$\exp\Big\{-\frac{\rho_e^{mm}}{2}(\beta_{e,m}-\hat{\mu}_e^m)^2\Big\}$$

$$\hat{\mu}_e^m=\mu_e^m+\textstyle\sum_{h=1,\neq m}^{M}(\beta_{e,h}-\mu_e^h)\rho_e^{hm} \tag{20-c}$$

where $\bar{\delta}_{ie}^k$ is a dummy variable, which is 1 when the prior condition state $d(\tau_A^k)=i$ of the inspection sample $k$ is equal to the prior condition state of the Gibbs sampling $e$ and 0 when they are different. $\mu_e^m$ is the $m$-th element of the prior expectation vector $\boldsymbol{\mu}_e$, and $\rho_e^{hm}$ is the $(h,m)$-th element of the prior variance-covariance matrix $\boldsymbol{\Sigma_e}^{-1}$. In addition, $\sum_{h=1,\neq m}^{M}$ means the sum of the elements from 1 to $M$, excluding $m$. It is possible to generate samples from these conditional probability density functions, and calculate statistics values regarding the posterior distribution of the parameters $\boldsymbol{\beta}$ using the samples. In short, the Gibbs sampling algorithm can be summarized in following steps:

Step 1: Specify the initial parameter $\beta(0) = (\beta_{1,1}(0),\cdots,\beta_{J-1,M}(0))$. Set $n=1$, and the number of samples $\overline{n}$.
Step 2: Generate $\beta(n) = (\beta_{1,1}(n),\cdots,\beta_{J-1,M}(n))$ as follows:
Generate $\beta_{1,1}$ randomly from $\pi(\beta_{1,1}|\beta^{-(1,1)}(n-1),\bar{\xi})$.
Generate $\beta_{1,2}$ randomly from $\pi(\beta_{1,2}|\beta^{-(1,2)}(n-1),\bar{\xi})$.
Generate $\beta_{e,m}$ randomly from $\pi(\beta_{e,m}|\beta^{-(e,m)}(n-1),\bar{\xi})$.
Generate $\beta_{J-1,M}$ randomly from $\pi(\beta_{J-1,M}|\beta^{-(J-1,M)}(n-1),\bar{\xi})$
Step 3: When $\underline{n}$ is sufficiently large and $n>\underline{n}$, record $\beta(n)$.
Step 4: When $n=N$, terminate the calculation. If $n<\overline{n}$, set $n=n+1$, and return to Step 2.

In the above Gibbs sampling, the transition kernel is defined as follows:

$$K(\boldsymbol{\beta}(n-1),\boldsymbol{\beta}(n)|\bar{\boldsymbol{\xi}})=\prod_{e=1}^{J-1}\prod_{m=1}^{M}\pi(\beta_{e,m}(n)|\boldsymbol{\beta}^{-(e,m)}(n-1),\bar{\boldsymbol{\xi}}) \tag{21}$$

where $\boldsymbol{\beta}(n)$ $(n=0,1,\cdots)$ is the Markov chain containing the transition kernel $K(\boldsymbol{\beta}(n-1),\ \boldsymbol{\beta}(n)|\bar{\boldsymbol{\xi}})$. In addition, the steady state of this Markov chain is represented by $\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}})$. Assuming that the Markov chain has reached the steady state while $\underline{n}$ is sufficiently large, it is possible to consider $\boldsymbol{\beta}(n)$ $(n=\underline{n}+1,\underline{n}+2,\cdots,\overline{n})$, which has been obtained in the Gibbs sampling, as the sample from the posterior probability density function $\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}})$. Accordingly, by using these samples, it is possible to calculate the statistics values regarding the posterior distribution of the parameter vector $\boldsymbol{\beta}$. In order to conduct the Gibbs sampling, it is necessary to obtain $(J-1)$ $M$ conditional posterior probability density functions $\pi(\beta_{e,m}|\boldsymbol{\beta}^{-(e,m)},\bar{\boldsymbol{\xi}})$ $(e=1,\cdots,J-1, m=1,\cdots,M)$. In general, adaptive rejection sampling (ARS) is effective as a method for obtaining samples from a probability density function $f(x)$, whose $\ln[f(x)]$ becomes a concave function (Gilks and Wild 1992). It is possible to verify that the conditional posterior probability density function $\pi(\beta_{e,m}|\boldsymbol{\beta}^{-(e,m)},\bar{\boldsymbol{\xi}})$ does not satisfy the above mentioned characteristics, but in many cases, $\ln[\pi(\beta_{e,m}|\boldsymbol{\beta}^{-(e,m)},\bar{\boldsymbol{\xi}})]$ is regarded as a logarithmic concave function, and so this study adopts the ARS as a method for sampling the parameters $\boldsymbol{\beta}$ of the posterior distribution from Eq. (20).

## 4.4 Setting of a Priori Distribution

When conducting the Bayesian estimation of the MUSTEM model, the prior information possessed by engineers can be expressed as a prior distribution. However, as mentioned above, there is no conjugate prior distribution in which the functional form of the prior distribution becomes the same as that of the posterior distribution. Therefore, it can be considered that the prior distribution based on prior experience does not represent the true probability density of unknown parameters. However, it is referred to as a tool for representing the past estimation results and the technical information possessed by engineers. In this study, the prior probability density function of the unknown parameters $\boldsymbol{\beta}_i$ $(i=1,\cdots,J-1)$ of the MUSTEM model is assumed to follow multidimensional normal distribution $\boldsymbol{\beta}_i\sim\mathcal{N}_M(\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i)$. By assuming such a relatively simple prior probability density function, it is possible to express the prior knowledge as the prior distribution in a simple form.

On the other hand, in the case of no information on the expectation or variance of the parameter vector $\boldsymbol{\beta}$, it is possible to use the following prior distribution whose variance is sufficiently large (Ibrahim et al. 2001):

$$\boldsymbol{\beta}_i\sim\mathcal{N}_M(\boldsymbol{O},\kappa_{\beta_i}\boldsymbol{I}) \tag{22}$$

where $\kappa_{\beta_i}$ is a sufficiently large positive number. In addition, $\boldsymbol{O}$ and $\boldsymbol{I}$ represent the zero vector and the unit matrix, respectively. It is also possible to use the following Jeffreys-type non-informative prior distribution

(Jeffreys 1961):

$$g(\beta_{i,m}) \propto \tilde{\beta}_{i,m} \tag{23}$$

where $\tilde{\beta}_{i,m}$ is an arbitrary constant that satisfies the condition: $-\infty < \tilde{\beta}_{i,m} < \infty$. Here, the Jeffreys-type non-informative prior distribution is called a non-holomorphic prior distribution, because it does not become 1 even when integrated in the domain of definition. A non-holomorphic prior distribution is not a probability density function, but it is possible to obtain a posterior distribution based on the Bayesian estimation rule, using a non-holomorphic priori distribution in Eq. (23) (Gamerman and Lopes 2006; Jeffreys 1961). That is, using the Jeffreys-type non-informative prior distribution, the conditional posterior probability density function in Gibbs sampling can be expressed as:

$$\begin{aligned}&\pi(\beta_{e,m}|\boldsymbol{\beta}^{-(e,m)},\bar{\boldsymbol{\xi}})\\ &\propto \prod_{i=1}^{e}\prod_{j=e}^{J}\prod_{k=1}^{K}\Big\{\exp\left(\beta_{e,m}x_m^k\right)^{\bar{\delta}_{ij}^k-\bar{\delta}_{ie}^k}\cdot\\ &\sum_{h=i}^{j}\prod_{l=i}^{h-1}\frac{1}{\theta_l^k-\theta_h^k}\prod_{l=h}^{j-1}\frac{1}{\theta_{l+1}^k-\theta_h^k}\exp(-\theta_h^k\bar{z}^k)\Big\}^{\bar{\delta}_{ij}^k}\end{aligned} \tag{24}$$

Eq. (24) shows the proportional relation, and so the Jeffreys-type non-informative parameter has been deleted from the right-hand side. As mentioned here, the conditional posterior probability density function used in Gibbs sampling varies according to the type of available prior information.

### 4.5 Statistics Values Regarding the Posterior Distribution

Based on samples obtained through the MCMC simulation, the statistical characteristics of parameter vector $\boldsymbol{\beta}$ can be analyzed. When the MCMC method is used, it is possible to express the posterior probability density function $\pi(\boldsymbol{\beta}|\boldsymbol{\xi})$ as an analytical function. Using the obtained samples, the distribution function and the density function are estimated in a non-parametric manner. Here, the samples obtained through the Gibbs sampling are represented by $\boldsymbol{\beta}(n)$ $(n = 1,\cdots,\overline{n})$, where $\boldsymbol{\beta}(n) = (\boldsymbol{\beta}_1(n),\cdots,\boldsymbol{\beta}_{J-1}(n))$. Among them, the first $\underline{n}$ samples are considered as the samples from the convergent process, and removed from the sample set. Then, the sample suffix set of parameters is defined as $\mathcal{M} = \{\underline{n}+1,\cdots,\overline{n}\}$. At this time, the joint probability density function $G(\boldsymbol{\beta})$ of the parameters $\boldsymbol{\beta}$ can be expressed by the following equation:

$$G(\boldsymbol{\beta}) = \frac{\#(\boldsymbol{\beta}(n) \le \boldsymbol{\beta}, n \in \mathcal{M})}{\overline{n}-\underline{n}} \tag{25}$$

where $\#(\boldsymbol{\beta}(n) \le \boldsymbol{\beta}, n \in \mathcal{M})$ represents the sum of samples satisfying the logical formula: $\boldsymbol{\beta}(n) \le \boldsymbol{\beta}, n \in \mathcal{M}$. In addition, the expectation vector $\tilde{\boldsymbol{\mu}}_i(\boldsymbol{\beta}_i)$ and the variance-covariance matrix $\tilde{\boldsymbol{\Sigma}}_i(\boldsymbol{\beta}_i)$ of the posterior distribution of the parameters $\boldsymbol{\beta}_i$ can be expressed by the following equations:

$$\begin{aligned}\tilde{\boldsymbol{\mu}}_i(\boldsymbol{\beta}_i) &= (\tilde{\mu}(\beta_{i,1}),\ldots,\tilde{\mu}(\beta_{i,M}))'\\ &= \Big(\sum_{n=\underline{n}+1}^{\overline{n}}\frac{\beta_{i,1}(n)}{\overline{n}-\underline{n}},\ldots,\\ &\sum_{n=\underline{n}+1}^{\overline{n}}\frac{\beta_{i,M}(n)}{\overline{n}-\underline{n}}\Big)'\end{aligned} \tag{26-a}$$

$$\tilde{\boldsymbol{\Sigma}}_i(\boldsymbol{\beta}_i) = \begin{pmatrix}\tilde{\sigma}^2(\beta_{i,1}) & \ldots & \tilde{\sigma}(\beta_{i,1}\beta_{i,M})\\ \vdots & \ddots & \vdots\\ \tilde{\sigma}(\beta_{i,M}\beta_{i,1}) & \ldots & \tilde{\sigma}^2(\beta_{i,M})\end{pmatrix} \tag{26-b}$$

where

$$\tilde{\sigma}^2(\beta_{i,m}) = \sum_{n=\underline{n}+1}^{\overline{n}}\frac{\{\beta_{i,m}(n)-\tilde{\mu}(\beta_{i,m})\}^2}{\overline{n}-\underline{n}} \tag{27-a}$$

$$\tilde{\sigma}(\beta_{i,m}\beta_{i,l}) = \sum_{n=\underline{n}+1}^{\overline{n}}\frac{\{\beta_{i,m}(n)-\tilde{\mu}(\beta_{i,m})\}}{\overline{n}-\underline{n}}\,\frac{\{\beta_{i,l}(n)-\tilde{\mu}(\beta_{i,l})\}}{\overline{n}-\underline{n}} \tag{27-b}$$

The $100(1-2\alpha)\%$ confidence interval can be defined as $\underline{\beta}_{i,m}^{\alpha} < \beta_{i,m} < \overline{\beta}_{i,m}^{\alpha}$, using the sample order statistic $(\underline{\beta}_{i,m}^{\alpha},\overline{\beta}_{i,m}^{\alpha})$ $(i = 1,\cdots,J-1,\quad m = 1,\cdots,M)$:

$$\underline{\beta}_{i,m}^{\alpha} = \arg\max_{\beta_{i,m}(n^*)}\left\{\frac{\#(\beta_{i,m}(n) \le \beta_{i,m}(n^*), n \in \mathcal{M})}{\overline{n}-\underline{n}} \le \alpha\right\} \tag{28-a}$$

$$\overline{\beta}_{i,m}^{\alpha} = \arg\min_{\beta_{i,m}(n^{**})}\left\{\frac{\#(\beta_{i,m}(n) \ge \beta_{i,m}(n^{**}), n \in \mathcal{M})}{\overline{n}-\underline{n}} \le \alpha\right\} \tag{28-b}$$

### 4.6 Bayesian Updating Rule

In the Bayesian estimation, if there exists a conjugate prior distribution while the prior and posterior distributions have the same function form, it is possible to update the Bayesian estimates of unknown parameters, using the newly obtained data. However, in the case of the MUSTEM model, the conjugate prior distribution does not exist. Therefore, when the Bayesian updating is conducted, it is necessary to accumulate all past monitoring data that are used for the model estimation.

Here, consider a case in which the posterior distribution in Eq. (18) regarding the unknown parameters of the Markov deterioration hazard model has been obtained, using the first monitoring data $\bar{\boldsymbol{\xi}}^0$. Then, discuss the problem of updating the posterior distribution of the unknown parameters, using the second monitoring data $\bar{\boldsymbol{\xi}}^1$. The available information regarding the inspection sample $k$ is $\boldsymbol{\xi}^k = (\bar{\delta}_{ij}^k, \bar{z}^k, \bar{\boldsymbol{x}}^k)$. In addition, the database pooling the first monitoring data

$\bar{\boldsymbol{\xi}}^0 = (\bar{\delta}_{ij}^k, \bar{z}^k, \bar{\boldsymbol{x}}^k)$ $(k = 1, \cdots, k_1)$ and the second monitoring data $\bar{\boldsymbol{\xi}}^1 = (\bar{\delta}_{ij}^k, \bar{z}^k, \bar{\boldsymbol{x}}^k)$ $(k = k_1 + 1, \cdots, k_1 + k_2)$ is defined. Assuming that the posterior probability density function of the unknown parameter vector in the first Bayesian estimation is $\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}^0)$, the posterior density function $\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}^0, \bar{\boldsymbol{\xi}}^1)$ of the unknown parameter vector after the Bayesian updating can be expressed as follows:

$$\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}^0, \bar{\boldsymbol{\xi}}^1) \propto \mathcal{L}(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}^1)\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}^0) \propto \mathcal{L}(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}^0, \bar{\boldsymbol{\xi}}^1) \prod_{i=1}^{J-1} g(\boldsymbol{\beta}_i|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (29)$$

where $\mathcal{L}(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}^0, \bar{\boldsymbol{\xi}}^1)$ is the likelihood function defined by using the database pooling the data of the first and second inspections. On the other hand, $g(\boldsymbol{\beta}_i|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ represents the prior distribution of $\boldsymbol{\beta}_i$ used in the first Bayesian estimation. Accordingly, the posterior distribution after the Bayesian updating becomes as follows:

$$\pi(\boldsymbol{\beta}|\bar{\boldsymbol{\xi}}^0, \bar{\boldsymbol{\xi}}^1) \propto \prod_{i=1}^{J-1} \prod_{j=i}^{J} \prod_{k=1}^{k_1+k_2} \Big\{ \prod_{l=i}^{j-1} \theta_l^k \cdot \sum_{h=i}^{j} \prod_{l=i}^{h-1} \frac{1}{\theta_l^k - \theta_h^k} \prod_{l=h}^{j-1} \frac{1}{\theta_{l+1}^k - \theta_h^k} \exp(-\theta_h^k \bar{z}^k) \Big\}^{\bar{\delta}_{ij}^k} \cdot \prod_{i=1}^{J-1} \exp\Big\{ -\frac{1}{2}(\boldsymbol{\beta}_i - \boldsymbol{\mu}_i)\boldsymbol{\Sigma_i}^{-1}(\boldsymbol{\beta}_i - \boldsymbol{\mu}_i)' \Big\} \quad (30)$$

That is, as obvious from Eq. (30), in order to update the posterior distribution of the unknown parameters, it is necessary to define a likelihood function for the database with new monitoring data, and obtain a posterior distribution by Gibbs sampling. Also when a non-informative prior distribution is used, it is possible to update the posterior distribution of the unknown parameters through Gibbs sampling, by defining a likelihood function for the database with new additional data.

## 5 EMPIRICAL STUDY

### 5.1 Outline

In order to discuss the effectiveness of the methodology proposed in this study, the Bayesian estimation of the MUSTEM model is attempted, using the monitoring data of bridges managed by N city. This section briefly describes the information of inspection and monitoring data of the bridges in N city. As a concrete target for the Bayesian estimation, this section focuses on RC slabs, on which loading is directly imposed and which are important members. Table 1 shows the evaluation criteria for the 7 condition states in the visual inspection of RC slabs. The amount of accumulated data in N city is enormous, and the number of inspection data obtained from the two successive timings amounts to 32,902. However, in reality, the database does not have such a plentiful amount of inspection samples in many cases. Then, the second database is produced by extracting samples from such database, and the effectiveness of the proposed Bayesian estimation is empirically examined.

**Table 1.** Notation of condition states

| Condition state | Physical meaning (RC slab) |
|---|---|
| 1 | Newly established state. There is no sign of deterioration. |
| 2 | Intermediate level between 1 and 3 |
| 3 | Water leakage is observed at part of the slab. (Unidirectional cracking accompanied by water leakage, and spotted water leakage at the edges) |
| 4 | Intermediate level between 3 and 5 |
| 5 | Water leakage is observed in over 75% of the slab area. Peeling and falling are observed at part of the slab. Free lime is observed along the flange on the girder. |
| 6 | Intermediate level between 5 and 7 |
| 7 | Serious peeling and free lime are observed. Slab part falling or its sign is observed. |

### 5.2 Effectiveness of the Gibbs Sampling

The exponential hazard function used for the Bayesian estimation is specified as follows:

$$\theta_i^k = \exp(\beta_{i,1} + \beta_{i,2}x_2^k + \beta_{i,3}x_3^k) \quad (i = 1, \cdots, 6; k = 1, \cdots, K) \quad (31)$$

For the explanatory variables in the above Eq. (31), the explanatory variable pair selected in Tsuda et al. (2006b) is adopted without any modification. For $x_2^k$ and $x_3^k$, the average traffic volume and the slab area of the RC slab $k$ are used. Vector of unknown parameter is denoted as $\boldsymbol{\beta}_i = (\beta_{i,1}, \beta_{i,2}, \beta_{i,3})$. In Eq. (31), in order to satisfy the constraint condition $\theta_i^k > 0$, the exponential function-based regression equation is employed.

Table 2 shows the results of the Bayesian estimation of the MUSTEM model (sample average of parameters), using the prior distribution setting the prior information as

$$\boldsymbol{\beta}_i \sim \mathcal{N}_3(\boldsymbol{O}, \boldsymbol{I}) \quad (32)$$

In addition, the $\mathcal{Z}$-score in Geweke statistic test (Geweke 1992) is also shown. For conducting Gibbs sampling, the number of samples for the Markov chain to reach the steady state is set as $\underline{n} = 3,000$, which is considered as burn-in samples. As shown in Table 2, the Geweke test statistic $\mathcal{Z}_{\beta_{i,m}}$ is less than 1.96, so it is obvious that the convergent hypothesis cannot be nullified with the 5% significance level. In the following calculations, suppose $\overline{n} = 13,000$, $\underline{n} = 3,000$ samples are removed as the samples from the process of converging to the posterior distribution, and the analysis is carried out by using the remaining 10,000 parameter samples.

**Table 2.** Result of the Bayesian estimation (using the original database)

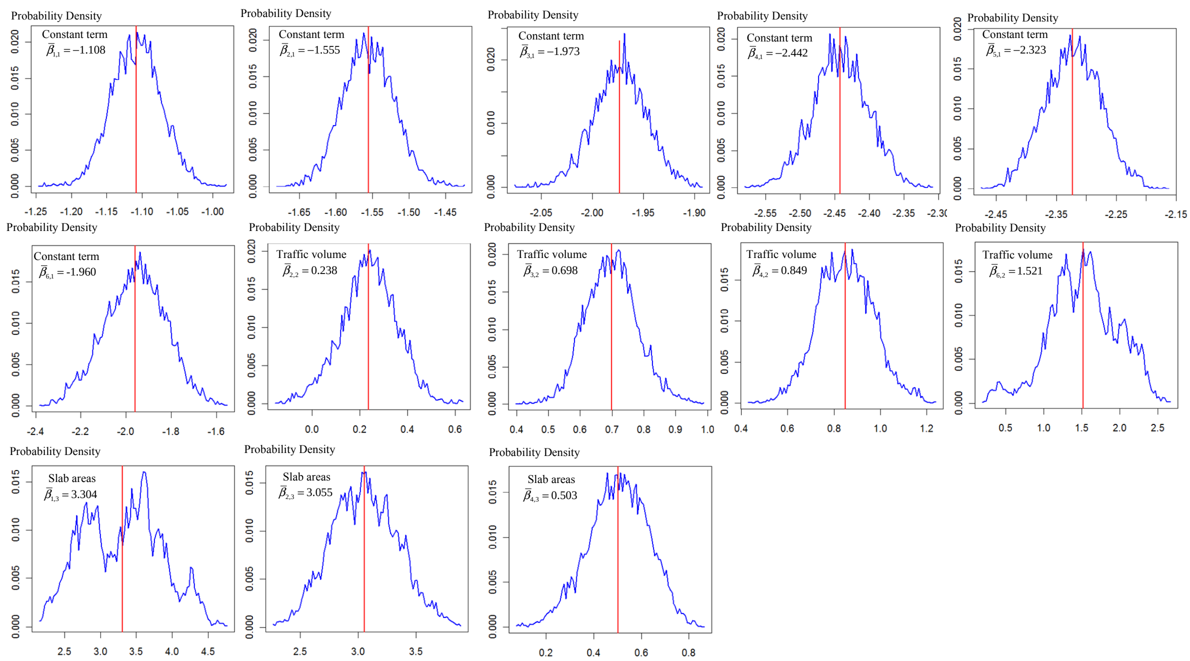
| Condition states | Prior distribution of Eq. 32 | | | Non-information prior distribution | | |
|---|---|---|---|---|---|---|
| | Constant term $\beta_{i,1}$ | TV $\beta_{i,2}$ | Slab area $\beta_{i,3}$ | Constant term $\beta_{i,1}$ | TV $\beta_{i,2}$ | Slab area $\beta_{i,3}$ |
| 1 | -1.108 | - | 3.304 | -1.100 | - | 2.854 |
| | (0.161) | - | (-0.572) | (0.160) | - | (0.160) |
| 2 | -1.555 | 0.238 | 3.055 | -1.555 | 0.237 | 3.033 |
| | (0.077) | (-0.004) | (-0.285) | (0.800) | (0.910) | (1.050) |
| 3 | -1.973 | 0.698 | - | -1.973 | 0.699 | - |
| | (0.116) | (0.143) | - | (0.270) | (1.240) | - |
| 4 | -2.442 | 0.849 | 0.503 | -2.422 | 0.849 | 0.506 |
| | (0.241) | (0.206) | (0.006) | (0.970) | (1.100) | (0.980) |
| 5 | -2.323 | - | - | -2.322 | - | - |
| | (0.060) | - | - | (0.210) | - | - |
| 6 | -1.960 | 1.521 | - | -1.955 | 1.538 | - |
| | (0.014) | (-0.016) | - | (0.030) | (0.090) | - |
| Log-likelihood | | -20062 | | | -20060 | |

Note: TV stands for average traffic volume.

The values of parameter shown in Table 2 are the mean values of parameter sample distributions, which are further illustrated in Figure 2. It can seen from the figure that the shapes of distributions for most of model's parameters are in Bell shapes and normal distributed, inferring high confidence on the estimation results.

For comparison, Table 3 also shows the estimation results of the model by applying the MLE method, using the original database. A greater detail of using the MLE method is explained in Tsuda et al. (2006b). In the case where the probability density function of the prior distribution regarding all parameters is positive and gentle in the domain, the posterior distribution approaches asymptotically to a normal distribution whose mean is the maximum likelihood estimate and variance is the inverse of Fisher Information.

As shown in Table 2, when the original database and the prior distribution of Eq. (32) are used, the sample mean is nearly equal to the maximum likelihood estimate shown in Table 3 for most parameters, and it is obvious that the Gibbs sampling is conducted efficiently. However, there exist parameters whose sample mean and maximum likelihood estimate are different from each other. This is considered due to the fact that the Gibbs sampling data is not enough to make




Note: $\overline{\beta}_{i,k}$ is the mean of 10,000 generated samples for $\beta_{i,k}$

**Figure 2.** Distribution of model's parameters $\boldsymbol{\beta}_i$.

the maximum likelihood estimate equal to the mean of the posterior distribution. One of the reasons why Gibbs sampling is inefficient in such as case that the variance of the prior distribution is small. Here, as shown in Table 2 and Table 3, there is little difference in log likelihood among the models.

Table 2 also shows the results of the Bayesian estimation with the non-informative prior distribution of Eq. (23). The non-informative prior distribution can be obtained by setting the variance of the prior distribution to be sufficiently large. The sample mean from the non-informative prior distribution is nearly equal to the maximum likelihood estimate, and there is no intrinsic difference between the Bayesian estimation of the MUSTEM model and the results of the estimation based on the MLE method.

**Table 3.** Estimation results based on the MLE method (Original database)

| Condition states | Constant term $\beta_{i,1}$ | Average traffic volume $\beta_{i,2}$ | Slab area $\beta_{i,3}$ |
|---|---|---|---|
| 1 | -1.101 | - | 2.888 |
| | (-30.35) | - | (2.76) |
| 2 | -1.555 | 0.239 | 3.029 |
| | (-42.44) | (1.90) | (9.16) |
| 3 | -1.973 | 0.696 | - |
| | (-68.45) | (7.33) | - |
| 4 | -2.440 | 0.845 | 0.513 |
| | (-57.38) | (6.55) | (3.85) |
| 5 | -2.323 | - | - |
| | (-50.51) | - | - |
| 6 | -1.951 | 1.544 | - |
| | (-14.81) | (3.56) | - |
| Log likelihood | | -20060 | |

Note: The numerical characters in the parentheses are t-values.

### 5.3 Data Accumulation Amount and Estimation Precision

*Deterioration Expectation Path*

In oder to verify the effectiveness and the precision of using our proposed methodology in different situations, we created a second database composing of 4 different group of data. The second database is created mainly based on the original data, which has been used with both MLE method and newly developed Bayesian estimation method (refer to Table 2). Following items detail the differences between 4 groups of data in the second database.

1. $D_{500}$ represents 500 history database obtained by random no-replacement sampling from the original database;
2. $D_{1000}$ denotes a total of 1,000 history database that is the sum of $D_{500}$ and the 500 new history database obtained by random no-replacement sampling from the original database;
3. Likewise, the $D_{1500}$ and $D_{2000}$ databases are produced.

Based on the 4 groups of data, the Bayesian updating of the MUSTEM model is attempted. The prior distribution is defined as follows, using the prior distribution in Eq. (22) whose variance is sufficiently large:

$$\boldsymbol{\beta} \sim \mathcal{N}_3(\boldsymbol{O}, 10000\boldsymbol{I}) \tag{33}$$

The procedures for the Bayesian updating of the MUSTEM model are summarized in following steps:

Step 1: A conditional posterior probability density function (20) is defined, using the initial prior distribution in Eq. (17) and the database $D_{500}$;
Step 2: Parameter samples are generated through Gibbs sampling from the conditional posterior probability density function (20);
Step 3: The mean, variance, and covariance are calculated from Eq. (26), using the generated parameter samples;
Step 4: A conditional posterior probability density function is defined, based on the likelihood function that has been defined, using the database $D_{1000}$ with new history data;
Step 5: Parameter samples are generated through Gibbs sampling, using the conditional posterior probability density function;
Step 6: The above procedures are repeated for each of $D_{500}$ to $D_{2000}$.

Table 4 shows the results of the Bayesian updating of the Markov deterioration hazard model for each of the databases $D_{500}$ to $D_{2000}$. One of the characteristics of the Bayesian estimation is that the probability distribution of the parameters of the model can be obtained. The confidence interval of the model parameter can be analyzed with the sample order statistics $(\underline{\beta}^{\alpha}_{i,m}, \overline{\beta}^{\alpha}_{i,m})$ $(i = 1, \cdots, 6; m = 1, 2, 3)$. Consequently, as shown in Table 4, as the amount of inspection data increases, the confidence interval becomes narrower. Furthermore, this section discusses how the precision in the estimation of the MUSTEM model changes through the Bayesian updating, using the deterioration expectation path. If the parameter value of the exponential hazard function is different, the deterioration expectation, which is derived from the different value, changes. Using the parameter sample $\boldsymbol{\beta}(n)$, which has been obtained through Gibbs sampling, the hazard rate of the Gibbs sample $n$ can be defined by means of following equation:

$$\theta^k_i(n) = \exp(\beta_{i,1}(n) + \beta_{i,2}(n)x^k_2 + \beta_{i,3}(n)x^k_3) \tag{34}$$

At this time, the expected duration $R^k_i(n)$ can be expressed by the following equation like Eq. (12):

$$R^k_i(n) = \frac{1}{\theta^k_i(n)} \tag{35}$$

In addition, the sample mean of the expected condition state duration can be calculated with the following equation:

$$E[R_i(n : \bar{\boldsymbol{x}}^k)] = \sum_{n=\underline{n}+1}^{\overline{n}} \frac{R_i(n : \bar{\boldsymbol{x}}^k)}{\overline{n} - \underline{n}} \tag{36}$$

**Table 4.** Results of the Bayesian updating of the MUSTEM model on 4 groups of data

| Parameters | $D_{500}$ | $D_{1000}$ | $D_{1500}$ | $D_{2000}$ |
|---|---|---|---|---|
| $\beta_{1,1}$ | -1.60(-2.26,-1.02 ) | -1.07(-1.45,-0.71) | -1.05(-1.35,-0.77) | -1.23(-1.51,-0.97) |
| $\beta_{1,3}$ | 21.5(0.16,49.1) | -0.69(-13.2,9.19) | -1.19(-13.5,8.64) | -2.05(-13.7,7.51) |
| $\beta_{2,1}$ | -1.92(-2.47,-1.42) | -1.65(-1.98,-1.33) | -1.70(-1.97,-1.44) | -1.66(-1.90,-1.43) |
| $\beta_{2,2}$ | 1.64(0.05,3.20) | 1.21(0.22,2.16) | 1.07(0.25,1.86) | 1.18(0.49,1.87) |
| $\beta_{2,3}$ | -5.52(-23.30,8.83) | -4.32(-14.00,3.44) | 1.72(-1.43,5.10) | 2.09(-0.06,4.31) |
| $\beta_{3,1}$ | -1.64(-1.97,-1.32) | -1.90(-2.14,-1.65) | -1.76(-1.96,-1.56) | -1.85(-2.02,-1.68) |
| $\beta_{3,2}$ | -0.30(-1.45,0.82) | 0.67(-0.14,1.47) | 0.51(-0.15,1.15) | 0.94(0.38,1.48) |
| $\beta_{4,1}$ | -2.13(-2.73,-1.57) | -2.30(-2.70,-1.92) | -2.27(-2.58,-1.97) | -2.31(-2.58,-2.05) |
| $\beta_{4,2}$ | -2.89(-5.58,-0.35) | -0.19(-1.52,1.14) | -0.14(-1.17,0.85) | -0.05(-0.91,0.81) |
| $\beta_{4,3}$ | 2.17(0.41,3.79) | 0.60(-0.84,1.85) | 0.89(-0.16,1.84) | 0.59(-0.49,1.52) |
| $\beta_{5,1}$ | -2.35(-3.01,-1.75) | -2.49(-2.99,-2.06) | -2.28(-2.65,-1.94) | -2.28(-2.60,-1.99) |
| $\beta_{6,1}$ | -4.92(-10.30,-1.62) | -1.86(-3.18,-0.77) | -2.50(-3.87,-1.36) | -1.79(-2.74,-0.98) |
| $\beta_{6,2}$ | 3.83(-5.33,14.60) | 0.24(-4.24,4.08) | 0.41(-4.28,4.77) | -1.22(-4.74,1.93) |

Note: The numerical characters in this table are the sample mean $\bar{\mu}(\beta_{i,m})$. The figures in the parentheses denote the 90% confidence region of each parameter.

**Table 5.** Sample mean of the mtp based on the Bayesian estimation method (Data set $D_{2000}$)

| Condition states | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Hazard rate |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.7649 | 0.2046 | 0.0285 | 0.0019 | 0.0000 | 0.0000 | 0.0000 | 0.2680 |
| 2 | 0 | 0.7618 | 0.2154 | 0.0220 | 0.0008 | 0.0000 | 0.0000 | 0.2720 |
| 3 | 0 | 0 | 0.8229 | 0.1682 | 0.0086 | 0.0003 | 0.0000 | 0.1949 |
| 4 | 0 | 0 | 0 | 0.9045 | 0.0907 | 0.0046 | 0.0002 | 0.1004 |
| 5 | 0 | 0 | 0 | 0 | 0.9032 | 0.0908 | 0.0060 | 0.1018 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0.8811 | 0.1189 | 0.1266 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | - |

**Table 6.** The mtp based on the MLE method (Original database - 32,902 data)

| Condition states | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Hazard rate |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.6862 | 0.2749 | 0.0368 | 0.0021 | 0.0001 | 0.0000 | 0.0000 | 0.3766 |
| 2 | 0 | 0.7756 | 0.2064 | 0.0174 | 0.0006 | 0.0000 | 0.0000 | 0.2541 |
| 3 | 0 | 0 | 0.8497 | 0.1422 | 0.0078 | 0.0002 | 0.0000 | 0.1628 |
| 4 | 0 | 0 | 0 | 0.8977 | 0.0974 | 0.0046 | 0.0003 | 0.1079 |
| 5 | 0 | 0 | 0 | 0 | 0.9067 | 0.0844 | 0.0089 | 0.0980 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0.8174 | 0.1826 | 0.2017 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | - |

Furthermore, in order to define the $100(1-2\alpha)\%$ confidence interval of the expected condition state duration, the sample order statistics $\underline{H}^{\alpha}(\bar{\boldsymbol{x}}^k), \overline{H}^{\alpha}(\bar{\boldsymbol{x}}^k)$ are defined as:

$$\underline{H}^{\alpha}(\bar{\boldsymbol{x}}^k) = \arg\max_{R_i^*(\bar{\boldsymbol{x}}^k)} \left\{ \frac{\#(R_i(n:\bar{\boldsymbol{x}}^k) \leq R_i^*(\bar{\boldsymbol{x}}^k), n \in \mathcal{M})}{\overline{n} - \underline{n}} \leq \alpha \right\} \quad \text{(37-a)}$$

$$\overline{H}^{\alpha}(\bar{\boldsymbol{x}}^k) = \arg\min_{R_i^{**}(\bar{\boldsymbol{x}}^k)} \left\{ \frac{\#(R_i(n:\bar{\boldsymbol{x}}^k) \geq R_i^{**}(\bar{\boldsymbol{x}}^k), n \in \mathcal{M})}{\overline{n} - \underline{n}} \leq \alpha \right\} \quad \text{(37-b)}$$

Then, the deterioration expectation path (hereinafter called "$100(1-2\alpha)\%$ confidence lower-limit curve") is produced based on the lower limit of the $100(1-2\alpha)\%$ confidence interval, and the deterioration expectation path (hereinafter called "$100(1-2\alpha)\%$ confidence upper-limit curve") is produced based on the upper limit. Figure 3 shows the $100(1-2\alpha)\%$ lower-limit curves and $100(1-2\alpha)\%$ upper-limit curves for $D_{500}$, $D_{2000}$. For the values characteristic variables $\bar{x}_2^k$ (traffic volume)) and $\bar{x}_3^k$ (slab areas), we normalized their values to be 0.2266 and 0.0431, respectively. The normalization was done by setting the maximum of traffic volume and slab areas to be equivalent to 1. In addition, $\alpha$ is set to be 0.05, so that the $100(1-2\alpha)\%$ confidence lower-limit curve and the $100(1-2\alpha)\%$ confidence upper-limit curve represent the 90% confidence lower-limit curve and the 90% confidence upper-limit curve, respectively.

As shown in Figure 3, it can be understood that as the accumulated data increases, the width between the 90% confidence lower-limit curve and the 90% confidence upper-limit curve becomes narrower. Figure 3 also shows the deterioration expectation path, which has been obtained from the MLE method. It is considered that as the amount of accumulated data increases,

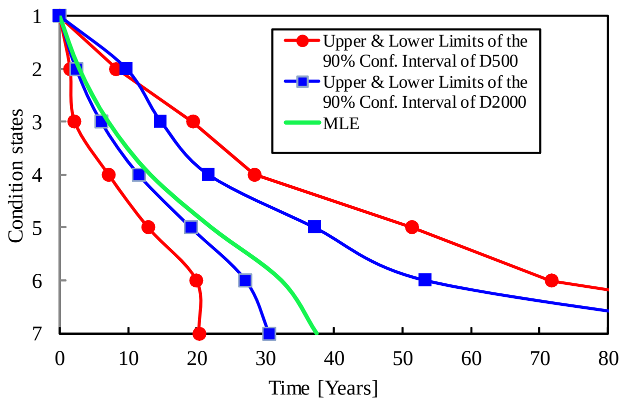



**Figure 3.** Deterioration expectation path.

the confidence lower-limit curve and the confidence upper-limit curve approach this deterioration expectation path.

### *Condition State Distribution*

In order to show the evolution of condition states over time, it is visually useful to present the curves of predicted deterioration. Moreover, it is also important to understand the time-series variation of the condition state distribution of the entire set of structures or the entire set of members, for carrying out asset management. Then, using the mtp, the time-series variation of the condition state distribution is analyzed.

If the hazard rate of the Gibbs sample $n$ defined by Eq. (34) based on the Bayesian estimation is obtained, the mtp can be estimated from Eqs.(7) and (9). Furthermore, as in the case of the expected condition state duration, the sample mean $E[p_{ij}(n : \bar{\boldsymbol{x}}^k)]$ and sample order statistic $(\underline{p}_{ij}^{\alpha}(\bar{\boldsymbol{x}}^k), \overline{p}_{ij}^{\alpha}(\bar{\boldsymbol{x}}^k))$ of the mtp of the inspection sample $k$ are defined. The definition of the concrete equation is omitted. It can be referred to Eqs. (36) and (37), if necessary. Table 5 shows the sample mean $E[p_{ij}(n : \bar{\boldsymbol{x}}^k)]$ of the mtp calculated from the results of the Bayesian estimation of $D_{2000}$, and Table 6 shows the mtp calculated from the maximum likelihood estimate. As for the mtp, it is possible to evaluate the confidence interval by calculating the sample order statistics.

Next, the condition state distribution is calculated. Consider the case in which all RC slabs in N city are new at timing $t$. The condition states are all 1, and so the initial value of the state vector is defined as $\boldsymbol{X}_t = (1, 0, 0, 0, 0, 0, 0)$. The arbitrary timing $t$ was specified as $t = 0$ and the calculation was repeated 100 times. Since the inspection interval of the estimated mtp is one year, by repeating the calculation 100 times, it is possible to show the evolution of condition states distribution over a span of 100 years.

Figure 4 shows the transition of the condition state distribution when the mtp (in Table 6) that has been estimated from $D_{2000}$ is used. There are few RC slabs whose condition state is still 1 at 20 years later. After that, deterioration progresses, and the occupancy rates

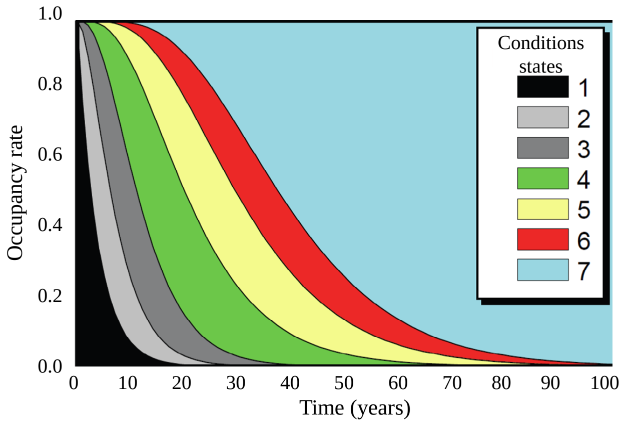



**Figure 4.** Condition state distribution, sample mean of the mtp $E[p_{ij}(n : \bar{\boldsymbol{x}}^k)]$.

of higher condition states (which indicate the progress of deterioration) increase. In addition, it can be understood that around 40 years after the start of deterioration, where the sample mean of the expected condition state duration is provided by the deterioration expectation path in the previous section, about 50% of RC slabs will likely to be in the condition state $i = 7$. In addition, based on the 90% confidence lower-limit and upper-limit values $(\underline{p}_{ij}^{\alpha}(\bar{\boldsymbol{x}}^k), \overline{p}_{ij}^{\alpha}(\bar{\boldsymbol{x}}^k))$ of the mtp, pessimistic and optimistic scenarios are also calculated respectively.

## 6 CONCLUSIONS

This paper has proposed a new methodology using Bayesian estimation method to estimate the parameter values of the MUSTEM model, which had been published earlier by the same authors. In previous paper, the author used MLE method to estimate model's parameter. The MLE method has been proved to be efficient in the case where more than 2000 data is used. However, it is not always possible in practice to collect a great number of data for use with the MLE method. In order to over come this problem, we develop a methodology using Bayesian estimation method to derive the model's parameters in the case of incomplete monitoring data. In the model, we assume the prior probability density function of model's parameter to follow multidimensional normal distribution, and then come up with the posterior probability density function of model's parameter and the complete likelihood function. To overcome the computational problem, we apply Bayesian estimation method, which is developed based on Gibbs sampling algorithm in the MCMC simulation. The proposed methodology has a preferable feature over the conventional MLE approach is that it offers a realistic way to utilize the experience-based information of experts in the field as prior information and combine it with actual monitoring data to predict the model's parameters.

To test the developed methodology in the real world, we carried out an empirical study on a same set of monitoring data that was used in our previously developed methodology using MLE approach. The monitoring data is visual inspection records on the evolution of condition states of RC slabs of bridge system in N city. The results indicate that when accurate prior information is provided, it is certain to obtain estimation results whose precision is at the same level as in MLE method.

Our proposed methodology has not discussed following points, which are considered for future extension of our research work.

1. The time-dependent deterioration process has not been discussed in the proposed Bayesian estimation method.
2. The measurement errors and bias in sampling, which are often embedded in monitoring data, have not been mentioned. For instance, under shortage of samples: when deterioration has progressed, some countermeasure is conducted, and so the amount of inspection samples regarding the condition states decreases.
3. The deterioration process concerning unobserved heterogeneity factor that inherits genuinely within individual structure or member of an infrastructure system.

## REFERENCES

Amemiya, T., and Boskin, M. (1974). "Regression analysis when the dependent variables is truncated lognormal, with an application to the determination of the duration of welfare dependency." *International Economic Review*, 15, 485.

Gamerman, D., and Lopes, H. F. (2006). *Markov Chain Monte Carlo, Stochastic Simulation for Bayesian Inference*. Chapman & Hall/CRC, Boca Raton, United States.

Geweke, J. (1992). *Bayesian Statistics 4*. Oxford University Press, Chapter Evaluating the Accuracy of Sampling-based Approaches to the Calculation of Posterior Moments.

Gilks, W., and Wild, P. (1992). "Adaptive rejection sampling for gibbs sampling." *Applied Statistics*, 41, 337–348.

Ibrahim, J., Ming-Hui, C., and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer-Verlag, New York, United States.

Jeff, G. (2006). *Bayesian Methods: A Social and Behavioral Sciences Approach*. Chapman and Hall/CRC, Boca Raton, United States.

Jeffreys, H. (1961). *Theory of Probability*. Oxford University Press, Oxford, United Kingdom.

Jido, M., Otazawa, T., and Kobayashi, K. (2008). "Optimal repair and inspection rules under uncertainty." *Journal of Infrastructure System*, ASCE, 14(2), 150–158.

Kaito, K., Abe, M., and Ishi-i, H. (2002). "Predicting the deterioration of structures based on actual inspection data." *Proceedings of IABSE symposium*, Melbourne, Astralia.

Kaito, K., Yasuda, K., Kobayashi, K., and Owada, K. (2005). "Optimal maintenance strategies of bridge components with an average cost minimizing principles." *Journal of Structural and Earthquake Eigineering*, JSCE, I-73(801), 83–96 (in Japanese).

Kobayashi, K., Kaito, K., and Lethanh, N. (2010). "Deterioration forecasting model with multistage weibull hazard functions." *Journal of Infrastructure Systems*, 16(4), 282–291.

Kobayashi, K., and Kuhn, K. (2007). "Decentralized life-cycle cost evaluation and aggregated efficiency." *The Management And Measurement Of Infrastructure: Performance, Efficiency and Innovation*, edited by C. Karlsson et al., Edward Elgar Press, Cheltenham, United Kingdom, 380–404.

Kobayashi, K., and Ueda, T. (2003). "Perspectives and research agendas of infrastructure managemenent." *Journal of Infrastructure Planning and Management*, JSCE, IV-61(744), 15–27 (in Japanese).

Lancaster, T. (1990). *The econometric analysis of transition data*. Cambridge University Press, Cambridge, United Kingdom.

Sugisaki, T., Kaito, K., and Kobayashi, K. (2006). "Statistical deterioration prediction considering nonuniformity of visual inspection cycle." *Journal of Structural Engineering*, JSCE, 52A(2), 781–790 (in Japanese).

Tobin, J. (1958). "Estimation of relationships for limited dependent variables." *Econometrica*, 26, 24–36.

Tsuda, T., Kaito, K., Yamamoto, H., and Kobayashi, K. (2006a). "Bayesian estimation of weibull hazard models for deterioration forecasting." *Journal of Construction Engineering and Management*, JSCE, 62(3), 473–491 (in Japanese).

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