The 2023 World Congress on
Advances in Structural Engineering and Mechanics (ASEM23)
GECE, Seoul, Korea, August 16-18, 2023

Reduction of Failure Risk using Local Evaluation Accuracy Improvement using GLMM

*Atsushi Iwasaki[1)]

[1)] *School of Science and Technology, Gunma University, 1-5-1 Tenjin-cho, Kiryu, Gunma 376-8515, Japan*
[1)] *aiwasaki@gunma-u.ac.jp*

ABSTRACT

This paper proposes a method for quantitative evaluation of the probability of failure and risk from condition monitoring diagnostic results using Bayes' theorem. When maintenance is performed using real-time monitoring results, ideally the diagnostic results should be free of inspection errors, but in general, most damage evaluation methods are subject to evaluation errors. The effect of evaluation error is not constant and depends on the damage's type, location, and degree of severity. Therefore, this study proposes a method to improve the accuracy of damage evaluation under specific conditions by using the random effect of GLMM. Two possible hazards arise from damage evaluation results: damage and unnecessary inspection implementation. Each is defined as accident risk and economic risk, respectively. The acceptable accident risk is considered to be constant. This paper clarifies the economic risk reduction effect of accuracy improvement of the local area by the proposed method.

1. INTRODUCTION

This research is concerned with a quantitative risk assessment method and a risk reduction method for maintenance using the damage identification results via inverse problem analysis. In recent years, the aging of many infrastructural facilities and the associated maintenance costs have become an issue. In many cases, external forces, material strength, internal damage, and other factors of infrastructure facilities are unknown, making it difficult to determine the timing of maintenance without being excessive safely. Condition monitoring is effective for appropriate maintenance. Regression analysis is often used for damage identification by inverse problem analysis based on monitoring. Regression analysis derives regression coefficients as constants by regression error minimization. It is easy to compute, but the output is a fixed value, and the average estimated value is the solution. However, failure occurs due to low-probability events and therefore requires consideration of the error distribution, which can occur if the estimation is significantly wrong. Therefore, the author's research group

---

[1)] Associate Professor

is studying a method for estimating the distribution of occurrence rates of damage parameters such as damage size and calculating the PoF(probability of failure) by deriving regression coefficients as a distribution using Bayesian estimation based on the damage identification results. In order to construct a method to make maintenance decisions quantitatively based on the estimation results, risk assessments are being conducted to quantitatively utilize the above(Iwasaki 2020). When maintenance is performed based on the estimation results using such a method, overestimating damage sufficiently smaller than the expected damage size naturally does not result in failure. If the damage is significant enough, a slight underestimation will not cause failure because the damage will be considered significant enough. In other words, it is important to improve the accuracy of estimating the range of certain specific damage levels to reduce the probability of failure. Therefore, this study investigates a method to reduce the PoF and risk by controlling the estimation accuracy in an arbitrary region using a generalized linear mixed model (GLMM), one of the hierarchical Bayesian models. The effectiveness of the proposed method in improving the accuracy of damage assessment in a specific region is verified, and an appropriate learning method of GLMM for risk reduction is discussed. In recent years, the aging of many infrastructural facilities and the associated maintenance.

2. QUANTITATIVE RISK ASSESSMENT OF RISK THROUGH MONITORING USING BAYESIAN ESTIMATION

2.1 Risk-Based maintenance for decision-making of maintenance activity

A method to make maintenance decisions based on quantitative or qualitative evaluation of risk is called risk-based maintenance(RBM). RBM is mainly used in the petroleum industry as a maintenance optimization method. API581 (USA)(API 2016) and Z107 (Japan)(HPI 2016) are significant examples of the RBM standard. When using RBM, risk indicators are plotted on a risk matrix. (Figure 1). The upper-right area is the high-risk area, and the lower-left area is the low-risk area. With RBM, maintenance programs are optimized by decreasing the equipment in a low- or high-risk area. Therefore, making decisions for maintenance programs, each combining various inspections, by calculating PoF from each inspection becomes possible because the consequence of failure depends on the equipment. Although RBM is well established as a method for time-based maintenance, its application as a condition-based maintenance method using ICT and other methods has been desired in recent years. This study is focused on a method for evaluating risk from the diagnostic result of real-time condition monitoring.

In this study, risk is defined as the product of an event's occurrence rate and the event's degree of impact.

Accident risk: Rate of failure occurrence × Impact of failure occurrence
Economic risk: Rate of wasteful inspections × inspection cost

Although the risk is divided into two categories, the degree of impact differs significantly due to accidents caused by underestimating severe damage and unnecessary inspections caused by overestimating small damage, respectively. This paper assumes that the acceptable risk of failure due to accidents is constant and examines a method to minimize the economic risk within a specific range of accident risk.

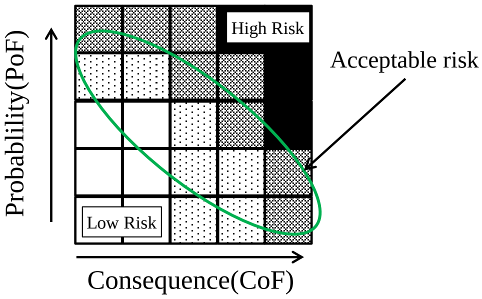


Fig. 1 Risk matrix for Risk Based Maintenance

2.2 PoF and risk evaluation by Bayesian estimation

Figure 2 shows the flow of PoF evaluation. Bayesian estimation is used to estimate the actual value incidence rate from the estimated value based on inverse problem analysis, and the PoF and risk evaluation are performed based on the results. First, the occurrence rate distribution of the estimated damage degree is estimated from the training data based on damage identification using GLMM as shown in Section 2.3. The occurrence rate distribution of estimated damage degree is the distribution of the estimated occurrence rate of damage degrees for the actual value of each damage degree (① in the figure). Next, the Bayesian method shown in the following equation is used to estimate the probability of occurrence of actual damage degrees for the results of degree identification by monitoring, i.e., the estimated damage degrees (Figure 3).

$$P(a_i \mid EstA_k) = \frac{P(a_i)P(EstA_k \mid a_i)}{\sum_j P(a_j)P(EstA_k \mid a_j)} \quad (1)$$

where $EstA_k$ is the estimated value of damage size a and a_i is the actual damage size. $P(a_i \mid EstA_k)$ is the posterior probability that the actual value of the damage is a_i when the estimated value $EstA_k$ is obtained. $P(a_i)$ is the prior probability, the probability that damage a_i will occur. The lower the degree of damage, the higher the probability of occurrence and an exponential distribution is assumed, as shown in the figure.

Then, the residual strength is estimated from the damage size and transformed into the residual strength distribution according to the damage mechanism that occurs (③ in the figure). Finally, since failure occurs when the external force exceeds the residual strength, the PoF for each estimated damage size is calculated by evaluating the PoF based on the limit state function method (④ in the figure). In this study, the external force distribution is set for the structure so that damage occurs on average at the damage size of 15 mm.

The overall risk evaluation is then performed for the estimated damage size obtained. Figure 3 shows the overall risk composition. The vertical axis is the product of the PoF, and the estimated damage occurrence rate PoO (adjusted loss probability, PoF_{adj}), and the horizontal axis is the estimated damage size. In this study, the structure was repaired if it was identified as above a certain damage level, and if it was identified as below a certain damage level, the structure was not repaired and continued to be used. Therefore, the accident risk is the sum of PoF_{adj} for damage size below the repair threshold, and the economic risk is the sum of the difference between PoO and PoF_{adj} for damage size above the threshold.

In this paper, the repair threshold is set so that the accident risk is constant (0.03), and the risk reduction effect is evaluated based on the increase or decrease of the economic risk.

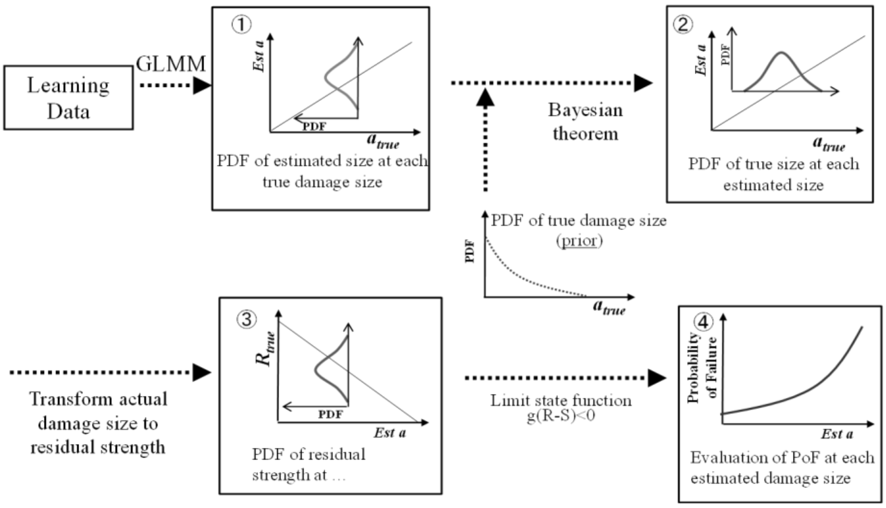


Fig. 2 Flow for the risk evaluation using GLMM

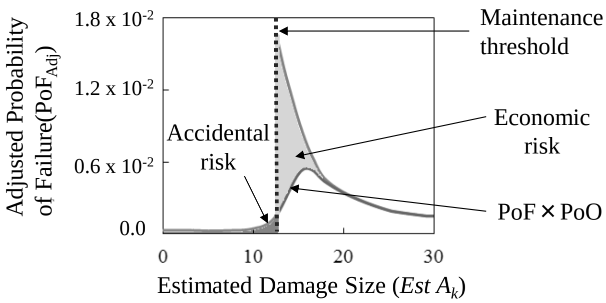


Fig. 3 Definition of accident risk and economic risk

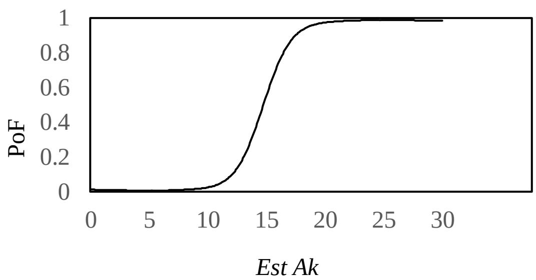


Fig. 4 PoF at each estimated damage size

2.3 Generalized Linear Mixed Model (GLMM)

Figure 4 shows the failure probabilities for the estimated damage size. The horizontal axis in the figure is the estimated damage size, and the vertical axis is the PoF. In this case, underestimation of damage of 10 mm or less and overestimation of damage of 20 mm or more do not significantly affect the PoF. On the other hand, a slight misevaluation of damage of about 15 mm can cause a significant change in the PoF. The proposed method attempts to reduce the damage probability by dividing the training data into multiple parts and using GLMM(Breslow 1993)(Iwasaki 2010)(Mcculloch 2008) for the regression method to improve the evaluation accuracy of specific parts.

The error structure must have normality in regression and multiple regression analysis, as typified by the least squares method. The generalized linear model is a regression model that uses the maximum likelihood method to obtain regression coefficients, which corresponds to the error structure of other distribution shapes. In addition, when the target of study, such as an organism, has individual differences, the effect of individual differences on the correlation between the explained variable and the explanatory variable becomes an error, which reduces the identification accuracy. The mixed model divides the correlation between the explained variable and the dependent variable into correlations that require identification (fixed effects) and correlations that do not require identification (random effects), aiming to improve the accuracy of fixed effects identification. The following equation expresses the model for a single regression of a GLMM model.

$$y = a + \delta_{1,i} + \left(b + \delta_{2,i}\right)x \qquad (2)$$

where x and y are the explanatory and explained variables, respectively; a and b are the intercept and slope of the fixed effects; and $\delta_{1,i}$ and $\delta_{2,i}$ are the respective random effects. i is the number of individuals, and individual differences are random effects. The number of degrees of freedom is (i +1)×2, which is vast if all the variate effects are estimated. In the GLMM, the random effects are not estimated for each individual, but only the standard deviations s_1 and s_2 of the random effects are estimated. In this case, the mean of random effects is assumed to be zero. The following equation shows the probability density function of random effect δ_1 when a normal distribution is assumed for the shape of the distribution of random effects.

$$g\left(\delta_{1,i} \mid s_1\right) = \frac{1}{\sqrt{2\pi}s_1} \exp\left\{-\frac{{\delta_{1,i}}^2}{2{s_1}^2}\right\} \qquad (3)$$

The probability that the explained variable takes y_i when the error structure is normally distributed and random effects are arbitrary $\delta_{1,j}$, $\delta_{1,k}$ is expressed as follows.

$$f\left(y_i \mid a, b, \sigma^2, \delta_{1,j}, \delta_{2,k}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{\left\{y_i - \alpha - \delta_{1,j} - \left(b + \delta_{2,k}\right)x_i\right\}^2}{2\sigma^2}\right\} \qquad (4)$$

where σ is the standard deviation of y. The likelihood L_i when the explained variable is y_i is defined from the integral of the respective random effects as follows.

$$L_i\left(a,b,\sigma^2,s_1,s_2 \mid y_i,x_i\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\left(y_i \mid a,b,\sigma^2,\delta_{1,j},\delta_{2,k}\right) g\left(\delta_{1,j} \mid s_1\right) g\left(\delta_{2,k} \mid s_2\right) d\delta_1 d\delta_2 \quad (5)$$

Therefore, the likelihood for all the observed data L is as follows.

$$L\left(a,b,\sigma^2,s_1,s_2 \mid \{y_i\},\{x_i\}\right)=\prod_{i=1}^{n} L_i\left(a,b,\sigma^2,s_1,s_2 \mid y_i,x_i\right) \quad (6)$$

where n is the number of observed data. In the case of multiple regression, the equation becomes more complex as the explanatory variables and random effects increase. The regression model is derived by performing maximum likelihood estimation and finding a, b, σ, s_1, and s_2 that maximize the log-likelihood. In this way, GLMM enables the calculation of fixed effects by deriving only the standard deviation of individual differences without deriving individual differences.

3. TARGET STRUCTURE AND ANALYSIS METHOD

3.1 Target structure and delamination identification using an electrical potential method

In this paper, the delamination identification of CFRP using the electrical potential method was conducted(Iwasaki 2005). Although CFRP has good mechanical properties in terms of specific stiffness and specific strength, its interlaminar strength is weak, which easily causes interlaminar delamination that is not visible from the outside due to weak impacts, resulting in significant degradation of compressive properties. In recent years, CFRP has been increasingly applied to the main structure of new large aircraft, and a simple evaluation method for the delamination of CFRP is desired. This study focuses on the electrical conductivity of carbon fibers embedded in CFRP as reinforcement fibers and investigates a method to identify the delamination size based on the change in electrical potential caused by the occurrence of delamination. Our group has conducted experiments using CFRP beam specimens with through-cracks to identify the location and size of delamination. It has been shown experimentally and analytically that the electrical potential method is effective and is also effective for delamination between buried actual layers. In order to implement these methods, it is essential to use an inverse problem method to link the measured electrical resistance change to the location and size of the delamination.

3.2 Analysis model

ANSYS was used for the analysis. The analysis was performed on a two-dimensional beam model of a $[0_2/90_2]_s$ orthogonal laminate shown in Figure 5. The longitudinal direction of the beam was taken as the 0º direction. Seven electrodes were placed in the 0º direction and the voltage variations at 6 locations between electrodes were analyzed. The elements were 4-contact quadrilateral elements with 0.25 and 0.05

mm in the 0º and 90º directions, respectively. The number of elements is 28160. The conductivity σ_i of each layer was determined experimentally for a volume fiber content of 0.62, with $\sigma_{90}/\sigma_0=3.7*10^{-2}$ and $\sigma_t/\sigma_0=3.8*10^{-2}$, respectively. The measurement is performed using a two-electrode method, and the amount of voltage change was derived from the voltage difference between two adjacent electrodes.

Delamination is modeled by eliminating conductivity between layers. The delamination was analyzed by changing the center position of the delamination by 5 mm from the test end to the opposite end, with sizes ranging from 3 to 29 [mm].

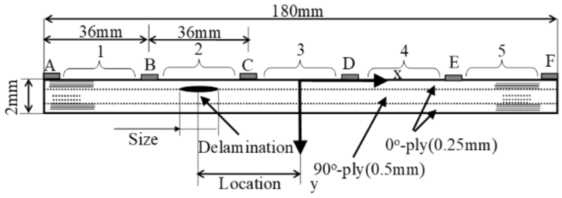


Fig. 5 Model of specimen for analysis

4. INVESTIGATION OF RANDOM EFFECTS EFFECTIVENESS FOR ECONOMIC RISK REDUCTION

4.1 Effects of division and subdivision of each damage region on damage evaluation accuracy

As shown in Figure 4, the effect of damage size evaluation error on the PoF varies significantly depending on the damage size. The damage size a of 0 to 30 [mm] was divided into three regions: the small damage region ($0\leq a<10$), the PoF rising region ($10\leq a <20$), and the severe damage region ($20\leq a$). In this section, the effect of the random effect of GLMM on damage identification is verified by subdividing the learning data in each region. The results are shown below. The following model was used to estimate the damage size.

$$y=\beta_0+\sum_{i=1}^{p}\left(\beta_i+\delta_{i,j}\right)X_i \qquad (7)$$

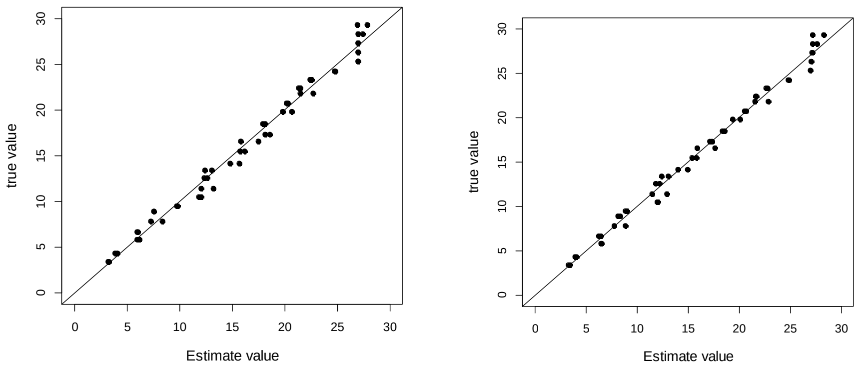
where y is the explained variable and X_i is the explanatory variable. p is the number of explanatory variables, β_i is the regression coefficient of the fixed effect and $\delta_{i,j}$ is the random effect on each explanatory variable. j is the number of divided regions. The explained variable is the size of the damage. The seven explanatory variables are the vector sum of the changes in the electrical potential between each electrode and the change in the electrical potential divided by the vector sum, as shown in the following equations.

$$X_7=\sqrt{\sum_{i=1}^{6}V_i^2} \qquad (8)$$

$$X_i=\frac{V_i}{X_7} \qquad (9)$$

Figures 6(a) and 6(b) show the regression results for the case of ordinary multiple regression and where the PoF rising region is divided. The horizontal axis shows the estimated values and the vertical axis shows the actual values. The figure shows that the GLMM qualitatively reduces the damage evaluation error, divided into several parts, especially in the PoF rising region.

For quantitative evaluation, Figure 7 shows the mean squared error when the small damage region, the PoF rising region, and both regions are divided into three parts, respectively. The horizontal axis of the figure shows the divided region, which is divided into six parts when both regions are divided. Without random effects indicate the results of ordinary multiple regression. The light dots indicate the mean squared error in the small damage region, the shaded area indicates the PoF rising region, and the gray area indicates the serious damage region. The GLMM reduces the error regardless of the region and improves the accuracy of the damage identification. This overall reduction in error results from improved accuracy due to the effect of hierarchization rather than an emphasis on specific regions. In addition, the mean squared error in the divided region is reduced significantly, indicating that the accuracy is improved by focusing on this region. When both regions are divided, the accuracy is improved in both regions, although the effect is lower than when each region is divided independently. In both cases, the accuracy of the divided regions was improved, and the use of GLMM effectively improved the accuracy of both the overall and local accuracy, which should be emphasized.



(a) without random effect (b) with random effect(PoF rising region)

Fig.6 Result of damage identification

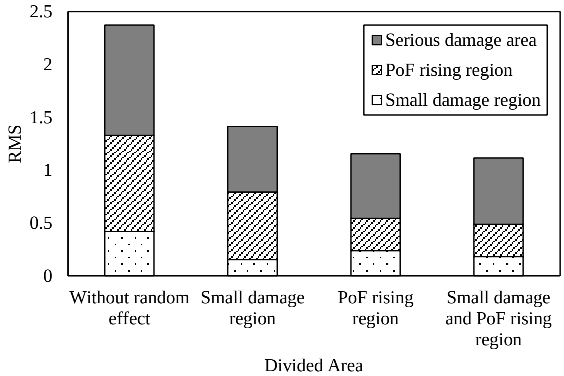


Fig.7 Result of damage identification

4.2 Effect of intercept random effect on evaluation accuracy

In section 4.1, random effects were given only to the slope of the regression model. This is because the rate of change between electrodes is zero when the damage size is zero. However, the PoF reising region data does not include data for small damage cases. Therefore, a random effect was also set for the intercept to test the effect. The regression model used is shown in the following equation.

$$y=\left(\beta_0+\delta_{0,j}\right)+\sum_{i=1}^{p}\left(\beta_i+\delta_{i,j}\right)X_i \qquad (10)$$

where $\delta_{0,j}$ is random effects on the intercept. Figure 8 shows the results for without random effects, intercept only, slope only, and both. The most improvement in accuracy is obtained when random effects are used for both the intercept and the slope. Although the effect of the intercept is smaller than that of the slope, the combination of the two can further improve the accuracy.

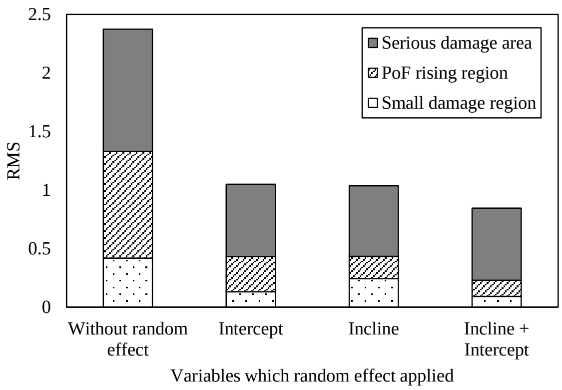


Fig. 8 Effect of the random effect on the intercept

4.3 Divided region s suitable for economic risk reduction

Figure 9 shows the risk reduction effect of improving the accuracy of the evaluation of local regions by using the random effect of GLMM. As shown in Section 2.2, the accident risk is assumed to be constant, and the economic risk is compared. The vertical axis of the figure shows the economic risk, and the horizontal axis shows the divided region. Random effects are given for both the intercept and the slope. As shown in the figure, economic risk reduction is achieved in all cases with random effects. The main reason for this can be attributed to the reduction of the overall evaluation error, as shown in Figure 6. Although the overall accuracy is lower when the random effect is set in the PoF rising region (Figure 6), the economic risk is lower when the random effect is set in the small damage region. Since the threshold was set so that the economic risk was constant, the accuracy of the left part in Figure 3 increased the threshold for repair execution. The accuracy in the PoF rising region is also considered adequate, but improving accuracy in the region of small damage is more effective in reducing economic risk.

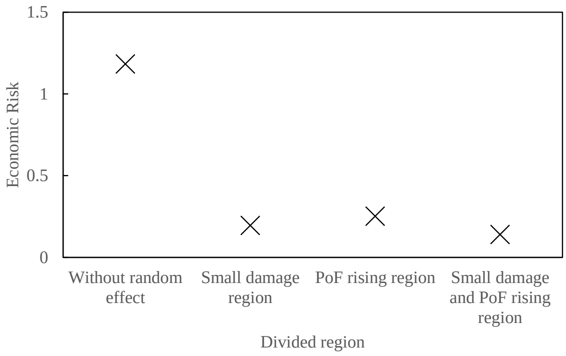


Fig. 9　Effect of the GLMM on risk reduction

5. CONCLUSION

This study investigated a method to reduce PoF and risk by controlling an arbitrary region's estimation accuracy using GLMM. The effectiveness of the proposed method in improving the accuracy of damage evaluation in specific regions was verified, and an appropriate GLMM learning method for risk reduction was investigated. The following is the conclusion of the study.

- The accuracy of the divided regions improved in both small damage regions and the PoF rising regions, indicating that it is possible to improve accuracy by dividing the regions for focused learning. In addition, when multiple regions are subdivided, the effect of subdivision is averaged. Therefore, it is adequate to focus division on necessary regions.
- The overall evaluation error is reduced by using GLMM, and the economic risk is lowered by proper division. Improving accuracy in the region of small damage is more effective in reducing economic risk than in the PoF rising region. The accuracy is the lowest when splitting is set for both regions. Therefore, improving the accuracy in the small-damage region is more effective in reducing economic risk, although improving the accuracy in the PoF rising region is also considered to be effective.

REFERENCES

American Petroleum Institute(2016), "API RP 581 Risk-Based Inspection Technology", API standards

Breslow, N.E., Clayton, D.G., (1993), “Approximate Inference in Generalized Linear Mixed Models”, J of the American Statistical Association, Vol.88, No.421, pp.9–25

High Pressure Institute of Japan(2016), "HPIS Z107 Risk Based Maintenance Handbook", HPI standard(in Japanese)

Iwasaki, A.(2020), "Bayesian Evaluation of Damage Risk from Monitoring Data", E-Journal of Advanced Maintenance, Vol.11, No.4, pp.147-152

Iwasaki, A., Todoroki, A., (2010), “Improvement of damage identification via inverse problem using the generalized linear mixed model”, J of JSME(A), Vol. 76, No. 768, pp. 32-39

Iwasaki, A., Todoroki, A., (2005), “Statistical Evaluation of modified Electrical Resistance Change Method for Delamination Monitoring of CFRP Plate, Structural health monitoring, An International Journal, Vol. 4, No. 2, pp. 119-136

Mcculloch, C.E., Searle, S.R., Neuhaus, J.M,(2008) “Generalized Linear, and Mixed Models”, Wiley Interscience