The 2016 World Congress on
The 2016 Structures Congress (Structures16)
Jeju Island, Korea, August 28-September 1, 2016

Bayesian approach for mixture modelling of stress response data

*Y.Q. Ni[1)] and R. Chen[2)]

[1), 2)] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong, China
[1)] ceyqni@polyu.edu.hk

ABSTRACT

Due to the multi-load effects, strain/stress response data acquired by structural health monitoring (SHM) system instrumented on a large-scale bridge often exhibits multimodality which cannot be adequately modelled by any standard unimodal probability distribution. This paper presents a finite mixture distribution model in the context of Bayesian inference, which is powerful for fitting the measured data with a heterogeneous population. Bayesian modelling approach rather than classical approach is developed for probabilistic parameter estimation of the mixture distribution model in viewing its capability of providing associated uncertainties of the estimates and handling the outlier data as well. The Markov chain Monte Carlo (MCMC) method is applied to evaluate the posterior densities. Synthetic data sets with multimodality are first used to verify the applicability and robustness of the Bayesian mixture distribution model. Later in the case study, the proposed method is applied to model the in-service stress response data of Tsing Ma Bridge under multi-load conditions.

Keywords: Bridge; SHM data; Strain/stress distribution; Multimodality; Bayesian mixture distribution model; MCMC method

1 INTRODUCTION

Application of long-term structural health monitoring (SHM) systems to large-scale bridges has attracted plenty of attention in the past decade (Ko & Ni, 2005; Brownjohn, 2007; Ni et al., 2011a). Through providing precise and real-time information about the structure and its environment, SHM system is aiming at assessing the structural condition with the help of data-driven diagnostic and prognostic tools. Extracting inherent features from comprehensive monitoring data would be the key work for the structural health and condition assessment while a variety of extraordinary stochastic phenomena exist which make a simply structured probability model inappropriate and

---
[1)] Professor
[2)] Ph.D. Student

unsatisfactory to catch the features. For example, it has been observed that the in-service stress response data acquired from hybrid bridges under the combination of highway, railway and wind loads exhibit multiple engendering effects, resulting in a heterogeneous data structure such that a standard distribution model is inadequate to characterize their statistical nature (Ni et al., 2011b; 2011c).

Among the existing statistical tools, the finite mixture model is a versatile and flexible method with which complex random phenomena can be favorably captured. It plays an important role in the inference of descriptive distribution models (McLachlan & Peel, 2004). Recently, the mixture model has been used in damage detection of aircraft wings (Qiu et al., 2014). Bayesian approach for parameter estimation of the mixture model has been addressed in the field of statistics (Diebolt & Robert, 1994; Lavine & West, 1992; Richardson & Green, 1997). Because of providing not only point estimators but also associated posterior distributions, Bayesian inference can better handle the inherent uncertainties of the model parameters. In this study, Bayesian approach for finite Gaussian mixture modelling of the probability distribution of heterogeneous strain/stress response data will be developed, along with the illustration of the proposed method with the use of real-world monitoring data acquired from an instrumented bridge.

2 BAYESIAN INFERENCE FOR GAUSSIAN MIXTURE MODEL

2.1 Finite Gaussian mixture model

The finite Gaussian mixture model can be expressed as

$$f(y_i;\boldsymbol{\theta}) = \sum_{j=1}^{J} \omega_j f_j(y_i;\mu_j,\sigma_j^2) \qquad i=1,\cdots,n;\, j=1,\cdots,J \tag{1}$$

$\boldsymbol{\Omega}=(\omega_1,\cdots,\omega_J)$ denotes a vector of mixing proportions or weights for J components and ω_j's are nonnegative quantities that sum to one, i.e. $0\le\omega_j\le 1, \sum\omega_j = 1$. For the Gaussian mixture model, the components are normal distributions which are expressed as $f_j(y_i) \sim N_j(\mu_j,\sigma_j^2)$.

In the framework of finite mixture distribution model, component indicator vector $\mathbf{Z}_i=(Z_{i1},\cdots,Z_{iJ})$ is introduced for each observation y_i, where Z_{ij} is defined to be one or zero depending on whether y_i in the mixture is equal to j or not

$$Z_{ij} = \begin{cases} 1 & \text{if the } i\text{th observation is drawn from the } j\text{th component} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Thus $\mathbf{Z}_i$ follows a multinomial distribution:

$$\mathbf{Z}_i \sim MN(1,\mathbf{\Omega}) \qquad \mathbf{\Omega} = (\omega_1,\cdots,\omega_J) \tag{3}$$

By including the component indicator vector $\mathbf{Z}_i$ as unknown parameters, the overall unknown parameters to be estimated in the Gaussian mixture model are

$$\mathbf{\theta} = \{\mathbf{\mu},\mathbf{\Sigma},\mathbf{\Omega}\} = \{\mu_1,\cdots,\mu_J,\sigma_1^2,\cdots,\sigma_J^2,\omega_1,\cdots,\omega_J\};\ \ \mathbf{Z} = \{\mathbf{Z}_1,\cdots,\mathbf{Z}_n\} \tag{4}$$

2.2 Prior selection

Bayes estimators for a finite mixture model can be well defined when the prior distributions are properly selected (McLachlan & Peel, 2004). In this study, we use conjugate priors for the parameters $\mathbf{\mu}$, $\mathbf{\Sigma}$, $\mathbf{\Omega}$ in the Gaussian mixture model. (μ_j,σ_j^2) are assumed to be mutually independent over the components. Normal-inverse-chi-squared prior is commonly used for the case that both mean and variance are unknown. A convenient parameterization is given as (Diebolt & Robert, 1994; Gelman et al., 2014)

$$\sigma_j^2 \sim InvC(\nu_j,s_j^2) \tag{5}$$

$$\mu_j \left| \sigma_j^2 \right. \sim N(\xi_j,\sigma_j^2/\kappa_j) \tag{6}$$

$InvC$ denotes the scaled inverse Chi-squared distribution, in which ν_j and s_j are degree of freedom and scale for σ_j, respectively. The conditional conjugate prior for μ_j is the normal distribution with mean ξ_j and variance σ_j^2/κ_j.

The prior distribution of $\mathbf{\Omega} = (\omega_1,\cdots,\omega_J)$ is assumed to be independent of μ_j and σ_j^2. A conjugate prior for $\mathbf{\Omega}$ is the Dirichlet distribution (McLachlan & Peel, 2004)

$$\mathbf{\Omega} \sim D(\alpha_1,\cdots,\alpha_J) \tag{7}$$

The relative sizes of the Dirichlet parameters α_j's describe the mean of the prior distribution of $\mathbf{\Omega}$, and the sum of α_j's is a measure of the strength of the prior distribution.

2.3 Posterior simulation using Gibbs sampler

Markov chain Monte Carlo (MCMC) algorithms can be used for posterior simulation and approximation when the joint posterior distributions have no analytical expression. Gibbs sampler is one of the frequently used MCMC algorithms and it is adopted here in the mixture model to estimate the unknown parameters.

Given the component indicator vector $\mathbf{Z}_i$, the conditional posterior distribution of

μ_j given σ_j^2 is a Gaussian distribution in the following expression

$$\begin{aligned}\mu_j\left|\sigma_j^2,\mathbf{y}_j,\mathbf{Z}_i\right. &\sim N(\xi_j^*,\sigma_j^2/\kappa_j^*)\\ &= N\left(\frac{\kappa_j\xi_j+n_j\bar{y}_j}{\kappa_j+n_j},\frac{\sigma_j^2}{\kappa_j+n_j}\right)\end{aligned} \tag{8}$$

$\mathbf{y}_j$ denotes the observations belonging to the jth component given $\mathbf{Z}_i$ while n_j is the number of observations. $\bar{y}_j$ is the sample mean of $\mathbf{y}_j$. The marginal posterior distribution for σ_j^2 is the scaled Inverse Chi-squared distribution:

$$\begin{aligned}\sigma_j^2\left|\mathbf{y}_j,\mathbf{Z}_i\right. &\sim InvC(v_j^*,s_j^{2*})\\ &= InvC\left(v_j+n_j,\frac{1}{v_j+n_j}\left(v_js_j^2+\sum_{i\in j}(y_i-\bar{y}_j)^2+\frac{\kappa_jn_j}{\kappa_j+n_j}(\bar{y}_j-\xi_j)^2\right)\right)\end{aligned} \tag{9}$$

v_j^*, s_j^{2*}, ξ_j^* and κ_j^* are the parameters of posterior distributions. From the algebraic forms of these parameters we can see that they are the combination of prior information and the information contained in the observations. The posterior distribution of the mixing proportion vector $\mathbf{\Omega}=(\omega_1,\cdots,\omega_J)$ is the Dirichlet distribution:

$$\mathbf{\Omega}\left|\mathbf{Z}_i\right. \sim D(\alpha_1+n_1,\cdots,\alpha_J+n_J) \tag{10}$$

Now we focus on the posterior distribution of the component indicator vector $\mathbf{Z}_i$ given the model parameters $\boldsymbol{\theta}$. From Eq. (3) we find that the posterior $\mathbf{Z}_i$ relies on the posterior $\mathbf{\Omega}$ which is denoted as $\boldsymbol{\tau}_i$ here. For the ith observation y_i, we have

$$\begin{gathered}\mathbf{Z}_i \sim MN(1,\boldsymbol{\tau}_i) \qquad \boldsymbol{\tau}_i=(\tau_1,\cdots,\tau_J)\\ \tau_j=\frac{f_j(y_i;\mu_j,\sigma_j^2)\omega_j}{\sum_{j=1}^{J}f_j(y_i;\mu_j,\sigma_j^2)\omega_j}\end{gathered} \tag{11}$$

in which τ_j means the posterior probability that observation y_i belongs to the jth component with y_i having been observed on it.

From Eqs. (8), (9), (10) and (11), we can summarize the procedures of Gibbs sampler for the Gaussian mixture model as illustrated in Fig. 1. Repeating the process, say $m=1,\cdots,M$, Gibbs sampler is run by simulating successively from the conditional distributions and replacing the conditioning parameters. After discarding a number of burn-in iterations, the generated random samples can be regarded as samples from the joint posterior distribution.

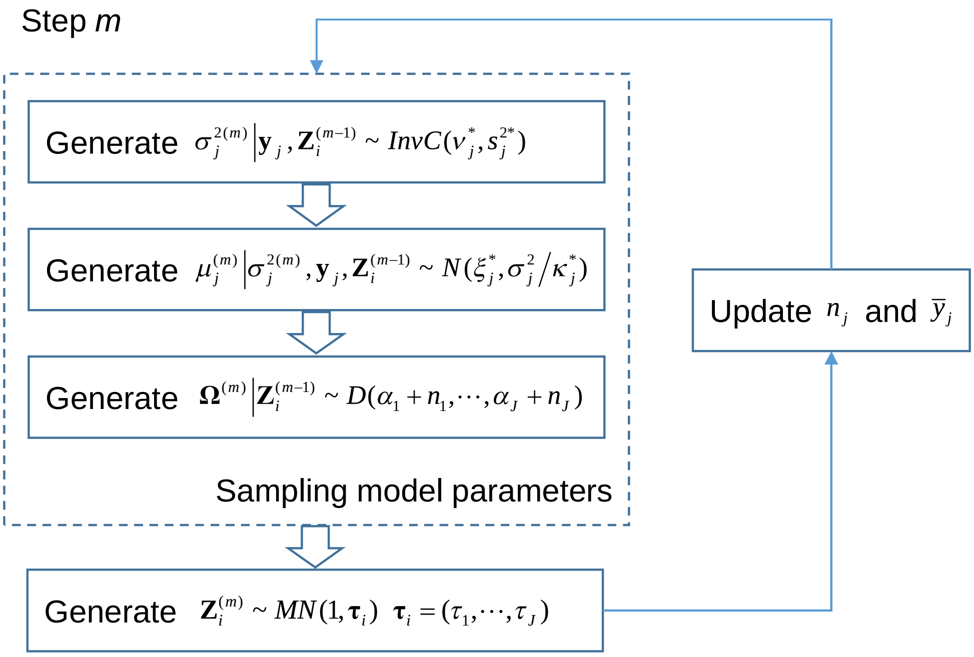

Fig. 1 Flowchart of Gibbs sampler for Gaussian mixture model

3 NUMERICAL EXAMPLES

The proposed algorithm is used to estimate the model parameters of a set of artificial samples, where the relative error between the estimators and true value can be obtained to evaluate the accuracy. We now consider a two-component Gaussian mixture model. Vague and diffuse prior specifications are preferred so that the estimation will be majorly influenced by the observed data. Herein, we set $v_j = 2$, $s_j^2 = \mathrm{var}(y) \times v_j$, $\xi_j = \mathrm{mean}(y)$, $\tau_j = 1$ and $\alpha_j = 5$ for $j = 1, 2$. Three cases are considered to evaluate the effectiveness of the Bayesian approach.

In Case 1, we consider the influence of different component means on the relative error of estimators, where all the parameters are assumed to be fixed except the 2nd component mean which is vary from 1.0 to 8.0 with the step of 1.0, namely Cases 1-1 to 1-8. In Case 2, the influence of different component variances on the relative error of estimators is considered, where the 2nd component variance is assumed to change from 1.0 to 6.0 with the step of 1.0 while the other parameters remain fixed (Cases 2-1 to 2-6). In Case 3, the influence of different weights assigned to each component on the relative error of estimators is further studied, where the 2nd component weight is assumed to change from 0.1 to 0.5 with the step of 0.1 while the 1st component weight from 0.9 to 0.5 respectively and the other parameters remain fixed (Cases 3-1 to 3-5).

The parameter configurations for all cases are illustrated in Tab. 1.

Tab. 1 Given parameter configurations

Case No.	Comp. No.	μ	σ^2	ω
1	1	1.0	1.0	0.5
	2	1.0 to 8.0*	1.0	0.5
2	1	1.0	1.0	0.5
	2	5.0	1.0 to 6.0**	0.5
3	1	1.0	1.0	0.9 to 0.5
	2	5.0	1.0	0.1 to 0.5***

Note: *Cases 1-1 to 1-8; **Cases 2-1 to 2-6; ***Cases 3-1 to 3-5

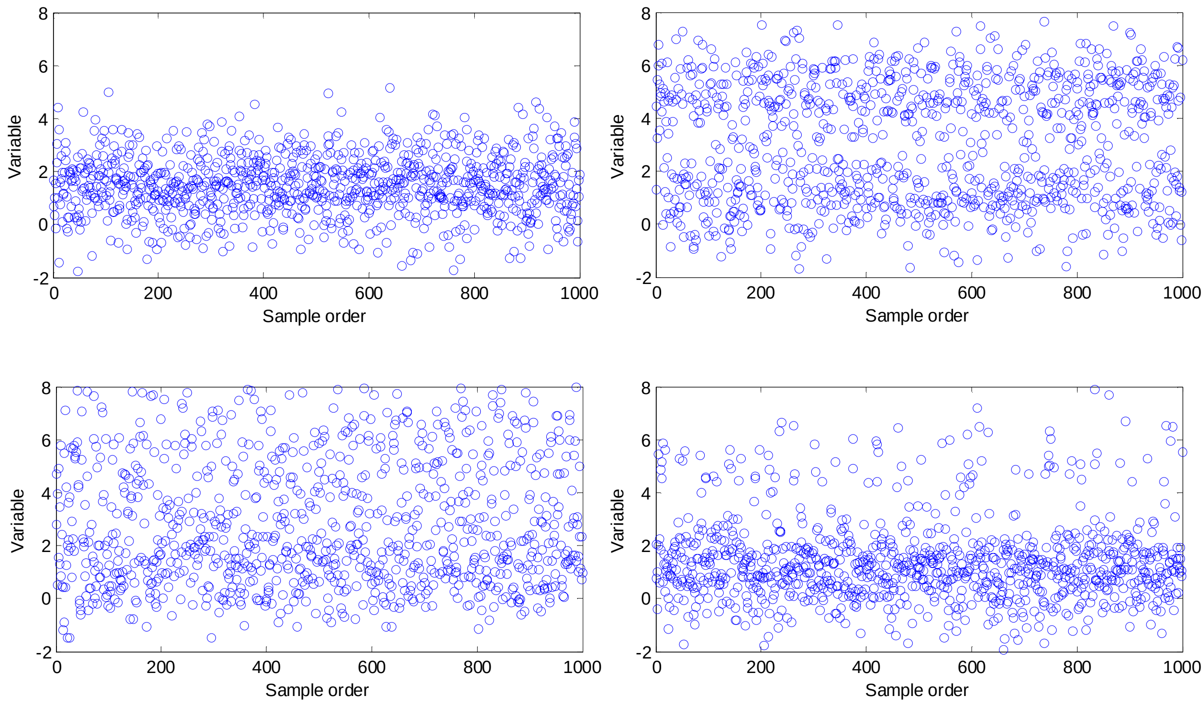

Fig. 2 Simulated random samples: a) Case 1-2; b) Case 1-5; c) Case 2-6; d) Case 3-1

For each case described in Tab. 1, 1000 random samples are generated as shown in Fig. 2, from which the number of Gibbs iterations and burn-in period are determined by visual inspection. The iteration paths as shown in Fig. 3 indicate that 10000 iterations and 2000 discards are sufficient for all cases. The estimation results of the model parameters are obtained as listed in Tab. 2. The relative errors in all cases are calculated and provided in Fig. 4.

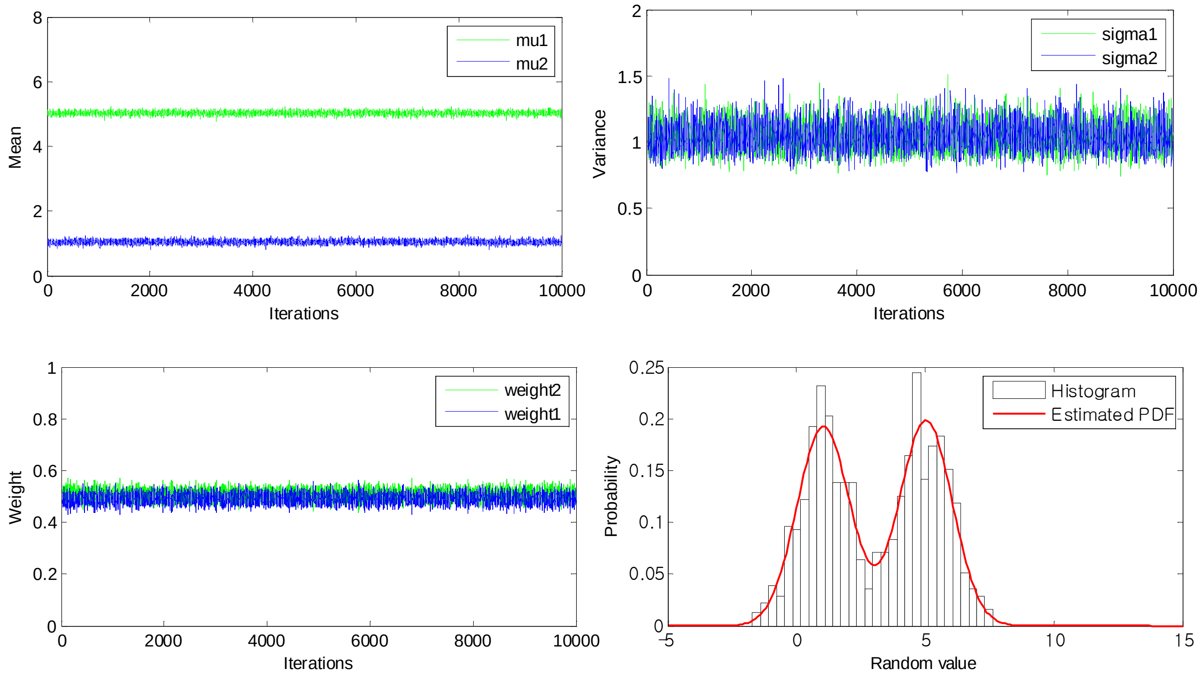

Fig. 3 Iteration paths for Case 1-5: a) iterations of component means; b) iterations of component variance; c) iterations of component weights; d) estimated PDF

Tab. 2 Estimation of parameters

Case	Comp. No.	μ	σ^2	ω
1-3	1	1.4276	1.1572	0.4992
	2	2.5693	1.1323	0.5008
1-5	1	1.0486	1.0459	0.4927
	2	5.0299	1.0416	0.5073
2-2	1	1.1092	1.0417	0.5302
	2	5.1238	1.8700	0.4698
2-4	1	1.0970	1.0345	0.5413
	2	5.4535	3.2686	0.4587
3-1	1	0.9932	0.9591	0.8927
	2	4.8908	1.4189	0.1073
3-3	1	1.0128	1.0081	0.6758
	2	4.9614	1.1409	0.3242

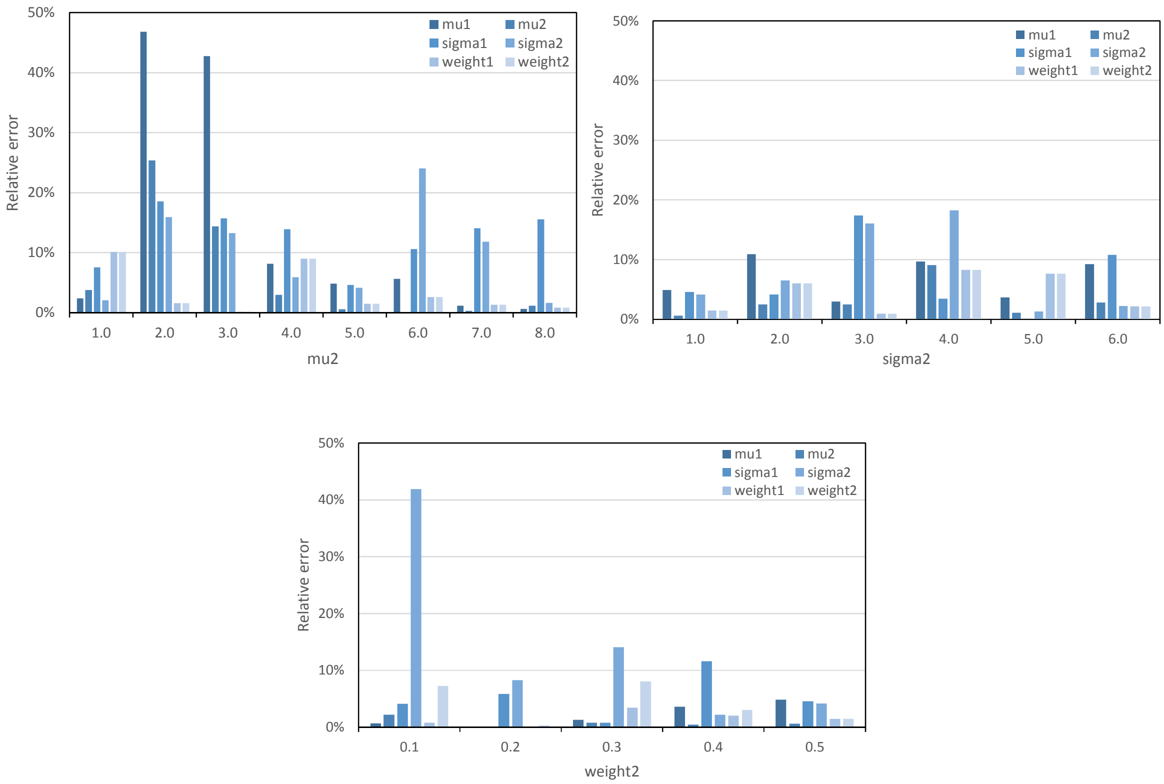

Fig. 4 Relative error: a) Cases 1-1 to 1-8; b) Cases 2-1 to 2-6; c) Cases 3-1 to 3-5

In Case 1, with the increase of the 2nd component mean (i.e. the distance between two component centroids becomes larger), much better estimators of the component means are obtained with the fact that the relative errors decrease dramatically. Meanwhile, the relative errors of the component variances fluctuate but the estimators of the component weights are of good quality. With the assumption that the component densities distribute separately (we set $\mu_1 = 1.0$, $\mu_2 = 5.0$), it is observed from Case 2 that good estimators of most of the parameters are still obtained with the increase of the 2nd component variance. The results of Case 3 reveals that different mixing proportions for components have little influence on the parameter estimation. From the above numerical examples, it can be inferred that when mixture samples exhibit obvious multimodality, the parameters in the Gaussian mixture model can be well identified by the proposed method.

4 APPLICATION TO STRESS RESPONSE DATA

4.1 Stress response of Tsing Ma Bridge

With the aid of the long-term SHM system instrumented on the Tsing Ma Bridge (TMB), evaluating the structural health and performance of the bridge becomes feasible through the analysis of huge amount of monitoring data. The monitoring of stain/stress responses of TMB due to various load combinations is essential since strain/stress response could provide intuitive verification of the structural performance and safety. A total of 110 strain gauges were installed on TMB aiming to measure dynamic strain responses at three bridge deck sections denoted by CH23488.00, CH23623.00 and CH24662.50 (chain mileages of the deck sections) as shown in Figs. 5 and 6 (Xia et al., 2012). The deployment locations of the strain gauges include the chord members (top chords, diagonal struts and bottom chords) of the longitudinal trusses, cross-frame chord members, bracing members, deck trough and rocker bearings at one tower. Three types of strain gauges, i.e. single, pair and rosette sensors, were used on the deck truss members. The data were continuously acquired at sampling rates of 25.6 and 51.2 Hz, respectively.

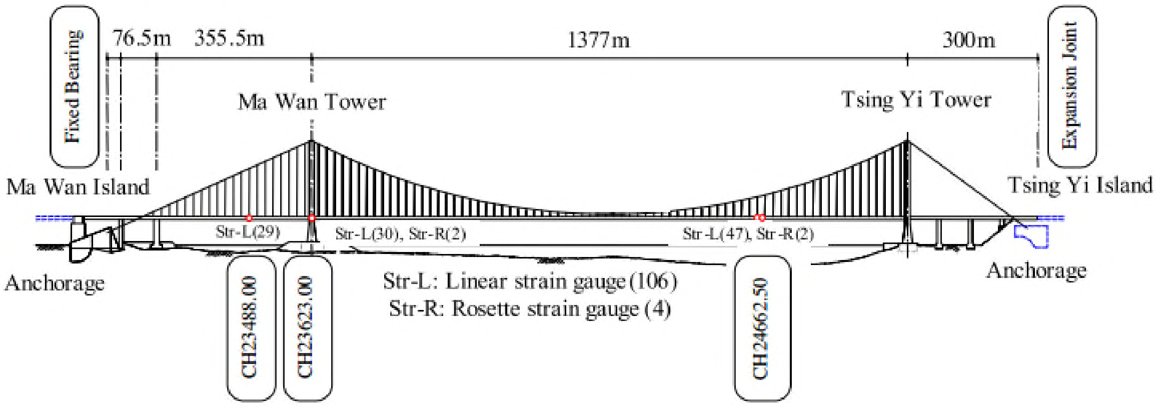

Fig. 5 TMB and strain monitoring sections

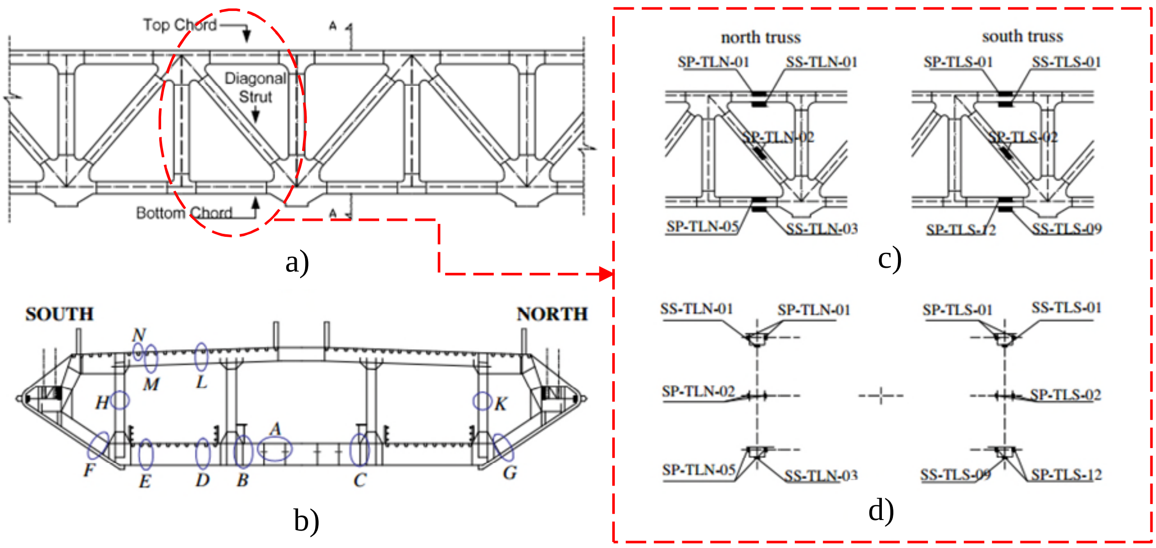

Fig. 6 Deck trusses of TMB and deployment of strain gauges

The in-service strain monitoring data acquired from the deck truss members of TMB are the result of combination of different loadings including vehicle, railway, monsoon, typhoon and temperature. The dead load caused initial strain cannot be measured since the strain gauges were installed after the completion of construction. Apart from the live loads, the temperature induced strain is considered to contribute little to the stress because it is released by the movement and rotation of the bridge deck at the expansion joint and bearings. A wavelet-based multi-component decomposition algorithm (Ni et al., 2011a) is used to eliminate the temperature induced strain aiming to obtain the live load induced strain which generates stress.

The strain monitoring data continuously collected during 24 hours at the bottom chords of the deck truss of Section CH24662.50 under highway traffic, railway traffic and normal wind loads are used in this study. Stress values are obtained by multiplying elastic modulus of steel by strain in viewing that the bridge is in elastic stage under operating environment. What we concern is those peak values in the stress time histories and they are extracted and grouped into peak stress sequences as shown in Fig. 7. This sample set contains 4200 peak values of stress response. From Fig. 7 we can observe that the peak stresses are randomly dispersed but centralized to two clusters, one is around 4.0 MPa and the other is around 10.0 MPa, exhibiting feature of group structure data. Outliers in the sample set are observed and they will be addressed in the proposed method. The histogram of peak stress shows distinct mixture property so that the mixture distribution model is desirable to characterize the multimodality of the peak stresses.

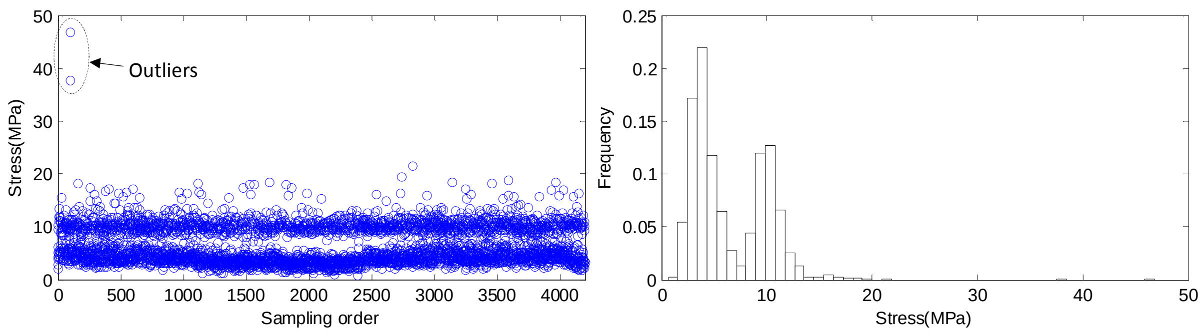

Fig. 7 Extracted peak stress sequences and stress histogram

4.2 Parameter estimation

The component number in the Gaussian mixture model should be first selected adequately otherwise the model will lose the essential ingredients contained in the stress information. A trial-and-error approach in conjunction with the Bayesian Information Criterion (BIC) is adopted to find the optimal component number. BIC is a measure that awards goodness of fit while penalizing model complexity (Kroese & Chan, 2014). Giving a set of competing models with different component numbers, the preferred model is the one with the maximum BIC value.

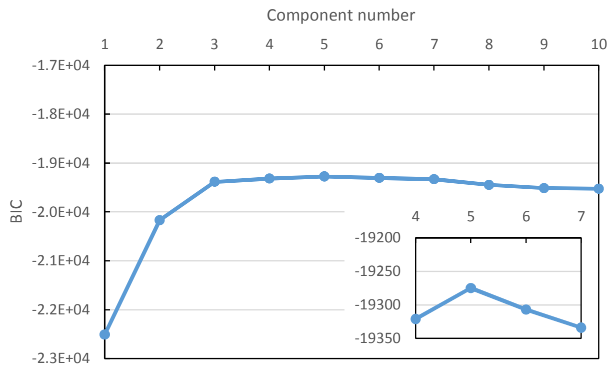

Fig. 8 Calculated BIC results

Fig. 8 shows the calculated BIC values among models with varying component number. It is observed that BIC gives a maximum in the case of five components so that the optimal component number for the Gaussian mixture model is determined as five. The posterior densities of the model parameters are obtained by applying the proposed Bayesian approach. Parameter estimation results including the sample mean and 95% credible interval (CI) are further obtained as listed in Tab. 3.

Tab. 3 Model parameter estimation

Comp.	μ		σ^2		ω	
No.	Mean	95% CI	Mean	95% CI	Mean	95% CI
1	12.3670	[11.3790, 13.3498]	9.8359	[7.2110, 13.0178]	0.0497	[0.0356, 0.0665]
2	3.6453	[3.5627, 3.7315]	0.9134	[0.8180, 1.0187]	0.4647	[0.4182, 0.5107]
3	10.0736	[10.0164, 10.1303]	1.1405	[1.0353, 1.2542]	0.3363	[0.3188, 0.3532]
4	18.5922	[11.9174, 28.6505]	96.7550	[51.6323, 170.9045]	0.0041	[0.0015, 0.0082]
5	5.0161	[4.7293, 5.3650]	2.1283	[1.7160, 2.6662]	0.1451	[0.1015, 0.1898]

It is found that the 2nd, 3rd and 5th components have almost represented the mixture property of the stress response by the fact that the sum of weights for these three components is very close to 1. It can be further interpreted that these three components represent the effect of superposition of highway, railway and wind loads. Weights of the 1st and 4th components are much smaller than others, implying that the observed stress responses have less probabilities to be clustered to these two components. These two components can be regarded as characterizing the statistical features of the outliers or other effects contained in the entire population.

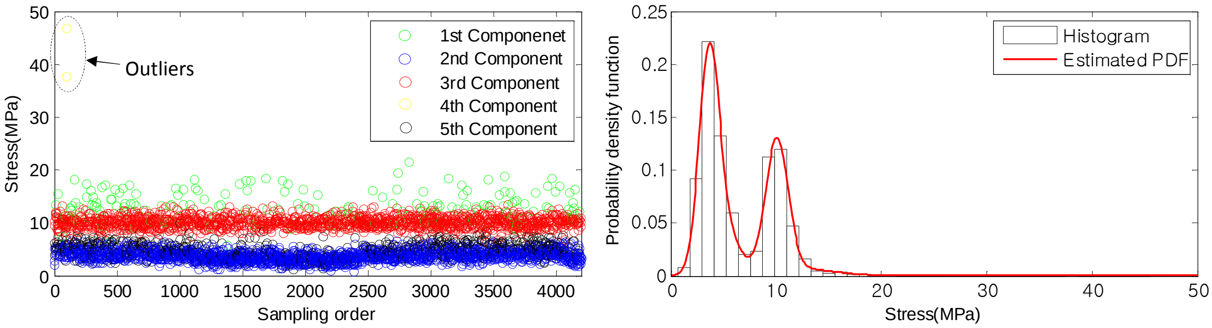

Fig. 9 Data cluster result and estimated PDF

By using the sample means of the posterior distributions of the model parameters, the estimated PDF for stress response under the combination of highway, railway and normal wind loads is constructed as depicted in Fig. 9. It is seen that the formulated Gaussian mixture model fits well with the histogram of stress response.

5 CONCLUSIONS

In recognition of the multimodal stress response data acquired from an instrumented bridge, a Bayesian Gaussian mixture modelling method along with prior specifications, posterior derivation and Gibbs sampling is presented in this paper. Numerical examples illustrate that parameter estimations of the Gaussian mixture model can be satisfactorily achieved by the proposed method in terms of good quality in results and robustness under different parameter conditions. The case study with the use of real-world data acquired from the Tsing Ma Bridge shows that the Gaussian mixture model comprising five component densities fits well with the measurement results.

ACKNOWLEDGEMENTS

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 152241/15E). The writers also wish to thank the Highways Department of Hong Kong SAR Government for providing the long-term structural health monitoring data of the Tsing Ma Bridge.

REFERENCES

Brownjohn, J.M.W. (2007), “Structural health monitoring of civil infrastructure.” *Philos. T. Roy. Soc. A.*, **365**(1851), 589-622.

Diebolt, J. and Robert, C. P. (1994), “Estimation of finite mixture distributions through Bayesian sampling.” *J. Roy. Stat. Soc. B. Met.*, **56**(2), 363-375.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2014), *Bayesian Data Analysis*, CRC Press, New York.

Ko, J. M. and Ni, Y. Q. (2005), “Technology developments in structural health monitoring of large-scale bridges.” *Eng. Struct.*, **27**(12), 1715-1725.

Kroese, D. P. and Chan, J. C. (2014), *Statistical Modeling and Computation*, Springer, New York.

Lavine, M. and West, M. (1992), “A Bayesian method for classification and discrimination.” *Can. J. Stat.*, **20**(4), 451-461.

McLachlan, G. and Peel, D. (2004), *Finite Mixture Models*, John Wiley & Sons, New York.

Ni, Y. Q., Wong, K. Y. and Xia, Y. (2011a), “Health checks through landmark bridges to sky-high structures”, *Adv. Struct. Eng.*, **14**(1), 103-119.

Ni, Y. Q., Xia, H. W., Wong, K. Y. and Ko, J. M. (2011b), “In-service condition assessment of bridge deck using long-term monitoring data of strain response.” *J. Bridge Eng.*, **17**(6), 876-885.

Ni, Y. Q., Ye, X. W. and Ko, J. M. (2011c), “Modeling of stress spectrum using long-term monitoring data and finite mixture distributions.” *J. Eng. Mech.*, **138**(2), 175-183.

Qiu, L., Yuan, S. F., Chang, F. K., Bao, Q. and Mei, H. F. (2014), “On-line updating Gaussian mixture model for aircraft wing spar damage evaluation under time-varying boundary condition.” *Smart. Mat. St.*, **23**(12), 1-14.

Richardson, S. and Green, P. J. (1997), “On Bayesian analysis of mixtures with an unknown number of components (with discussion).” *J. Roy. Stat. Soc. B. Met.*, **59**(4), 731-792.

Xia, H. W., Ni, Y. Q., Wong, K. Y. and Ko, J. M. (2012), “Reliability-based condition assessment of in-service bridges using mixture distribution models.” *Comput. Struct.*, **106**, 204-213.