

# Bayesian Model Calibration and Model Selection of Non-destructive Testing Instruments



Sharvil A. Faroz, Devang B. Lad and Siddhartha Ghosh

**Abstract** Concrete acquires a major share of infrastructure and building stock. However, degradation and deterioration of reinforced concrete (RC) structures have been a major concern for the construction industry in recent years. Evaluating the current health of the structure is important for taking a decision regarding future action on the structure. Structural health monitoring (SHM) becomes an essential step for an engineer to gain knowledge about the health of the structure. SHM faces various challenges due to site conditions as well as the limitations present in the NDT tool itself. SHM comprises of collecting health data using NDT tools and analyzing it with the physical or empirical model. These models are fitted using various techniques to develop the relationship between the NDT readings and the corresponding actual quantity of interest. Before the NDT tool is taken on site for actual investigation, the model should be calibrated in the laboratory. The model calibration of the NDT tool is prone to measurement uncertainties which are not properly incorporated in the commonly adopted regression method of calibration. This paper focuses on the model calibration and selection of the best model using Bayesian inference. Bayesian inference helps to quantify the measurement uncertainties. For this, a measurement error model (MEM) is adopted to relate the NDT readings to the property being estimated. An illustration of the calibration and selection process is demonstrated for the proposed approach. For the demonstration, we adopt rebound hammer which is one of the most common NDT tools used to evaluate the present strength of concrete by relating the NDT readings to the crushing strength values obtained in the laboratory. As multiple models are available in the literature for both the cases, Bayesian model selection method is used for selecting the most plausible

---

S. A. Faroz

Infrastructural Risk Management, Mumbai, India

e-mail: [iamsharvil.faroz@gmail.com](mailto:iamsharvil.faroz@gmail.com)

D. B. Lad · S. Ghosh (✉)

Structural Safety, Risk and Reliability Lab, Department of

Civil Engineering, Indian Institute of Technology Bombay, Mumbai, India

e-mail: [sghosh@civil.iitb.ac.in](mailto:sghosh@civil.iitb.ac.in)

D. B. Lad

e-mail: [devangladb@gmail.com](mailto:devangladb@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020

P. V. Varde et al. (eds.), *Reliability, Safety and Hazard Assessment*

*for Risk-Based Technologies*, Lecture Notes in Mechanical Engineering,

[https://doi.org/10.1007/978-981-13-9008-1\\_34](https://doi.org/10.1007/978-981-13-9008-1_34)

model from all available models. This will help us to identify which model represents the NDT instrument in the best possible way.

**Keywords** Bayesian inference · Measurement uncertainty · Model calibration · Model selection · NDT · RC structures · Rebound hammer · SHM

## 1 Introduction

With the commencement of the twentieth century, there was a surge in the use of the reinforced concrete structures. It was eventually when concrete structures started failing and deteriorating in an aggressive environment, engineers realized that concrete durability concerns were far more complex [1, 2]. Concrete degradation over an extended period with an increased rate of aggressive environments has challenged engineers for its maintainability and safe functioning. Infrastructure location exposes concrete to environment highly susceptible to degradation. Thus evaluation of its present state to closely accurate results is subject that to be dealt with caution. These factors lead to either repair, retrofit, rehabilitate, and replace the structure based on the vulnerability of the structure [3]. Before it is repaired or retrofitted, it is necessary that the condition of the present structure is evaluated. Several methods like destructive, semi-destructive, and non-destructive techniques are available for evaluation of concrete properties [1]. Spatial variation and instrument precision affect the engineering judgments. Various standards viz., [4–7] etc. are available which provides guidelines for the assessing concrete properties using a Non-Destructive Testing (NDT) method.

NDT inspection involves the measurement recorded on the structure, evaluation of concrete strength/property and analyzing the results obtained. This measurement is prone to various errors and contains inherent deficiency due to various factors involved in the inspection process [8]. The quality of measurement recorded depends on the calibration of the instrument before it is actually taken on the site. For example, rebound hammer before used on the actual structure, correlation graph is provided on the instrument that is calibrated using the laboratory data of concrete cubes [9]. Thus, it becomes necessary to calibrate the instrument correctly to incorporate the uncertainties involved in the process. Various techniques are available for calibrating the NDT instruments such as regression analysis, curve fitting, artificial neural networks (ANN), and Bayesian updating. A huge amount of work is carried out by researchers for calibration of NDT instruments and providing conversion models and calibration curves. For e.g., Ploix et al. [10] carried out bilinear regression between UPV versus water saturation and porosity rate where they successfully combined providing a surface plot, which shows that the properties of the concrete can be correlated while Sbartaï et al. [11] correlated UPV, RN, and compressive strength of concrete cubes using 3-D plots. EN 13791: 2007 [12] suggest the calibration of NDT instrument before using it for evaluation of structure. In case of corrosion, various models for Linear Polarisation Resistance (LPR) are developed based on time-varying nature

of the corrosion rate which is based on power law, Faraday’s law, Fick’s law, etc. [13]. These model parameters are to be first calibrated using experimental data to represent the actual phenomena in the most accurate way. The process of evaluating these parameters is called as model calibration. Once calibration of models is achieved, it is necessary to check the appropriateness of each model and select the best model which fits the calibration data with least error, and it is called as model selection [14, 15]. In this paper, a demonstration of the model calibration and model selection are discussed using Bayesian updating. This paper focuses on calibration through Bayesian updating of the instrument using a probabilistic measurement error model (MEM), in order to capture the measurement uncertainty and selecting the best model [8]. Major advantages of the Bayesian model calibration are that it allows to incorporate the variation in the calibration data and the parameters can be continuously updated with new incoming data. But however it does not take into effect the variation (or uncertainty) in the structure itself and does not involve special cases like carbonated concrete, corroded, etc.

## 2 Bayesian Updating

Bayes’ theorem is given by Eq. (1),

$$f(\theta|D^c) = \frac{f(D^c|\theta) \times f(\theta)}{\int_{\Theta} f(D^c|\theta)f(\theta)d\theta} \tag{1}$$

where  $f(\theta|D^c)$  is posterior probability,  $f(D^c|\theta)$  is the likelihood obtained from observed/measured/calibration data of a system, and  $f(\theta)$  is the prior distribution of the model parameters [8]. The prior and the likelihood are the pillars of Bayesian inference. The prior distribution of the parameters is obtained from the literature which is known as informative prior while if this is unknown a (non-informative) suitable distribution can be also be assumed. Likelihood function provides with the necessary input to update the past knowledge (prior) which is incorporated using experimental data. Generally, for  $n$  observed/measured statistically independent data, the likelihood function is given by [16],

$$f(D^c|\theta_l) = L(\theta_l) = \prod_{i=1}^n L_i(\theta_l) \tag{2}$$

Here, the subscript is added for  $\theta_l$  which corresponds to the specific model parameters. There are several limitations in the computation of posterior distribution for various reasons viz. (1) multiple number of parameters are involved, (2) large number of samples are required for simple Monte Carlo Simulation (MCS), (3) posterior is not of a standard form, etc. [8]. Hence in this study, we adopt Markov Chain Monte Carlo (MCMC) simulation where correlated random samples are generated

enabling the user with a less computational intensive for posterior distribution. Several MCMC techniques are available for posterior computation, this study adopts Metropolis–Hastings algorithm for Bayesian updating [17].

### 3 Bayesian Model Calibration

The calibration problem can be summarized as one relating the NDT measurement recorded to the quantity of interest defined using a mathematical model under some specific condition. The parameters of the mathematical model are to be evaluated and the quality of the fit obtained depends on the scatter of the data, calibration procedure adopted, number of training data (calibration data) available, etc. This may include uncertainties due to the measurement process (instrument) and model itself. These uncertainties present in the instruments can be accounted for by formulating a probabilistic measurement error model (MEM) to relate NDT readings to the true values. From the available deterministic models (Sect. 3.2), we may apply probabilistic context taking measurement uncertainty into account by adding a random variable  $E \sim N(\mu = 0, \sigma = e_0)$ .  $E$  is a correcting random variable where  $e_0$  is the random error/noise which is independent and identically distributed (i.i.d.) which represents measurement uncertainty [8]. The model now can be written as,

$$Y = M(B; X) + E \quad (3)$$

On the basis of the probabilistic analysis, here  $B$  of  $M(\cdot)$  are the parameters of the models that are needed to be evaluated. Such regression models involving inspection or measurement are MEM models.

Calibration is an inverse problem where the aim is to evaluate the values of parameters with the known value of  $Y$  and  $X$  (measurements which are known as training data). In Bayesian calibration, these parameters are modeled as random variables quantified by their respective probability distribution function (PDFs). The data obtained through repeated experiments are combined with the prior distribution of the parameters to obtain an updated posterior distribution. The prior can be represented by the joint PDF assuming that no correlation exists between each of the parameters,  $f(\theta_l|M_l)$  [18]. The likelihood function  $f(D^c|\theta_l, M_l)$  can be formulated as given in Eq. (2) where  $D^c$  is the training data and  $\theta_l$  are the model parameters for the particular model class,  $M_l$ . Hence from the Eq. (1), posterior is the product of prior and likelihood, given by [8],

$$f(\theta_l|D^c, M_l) = \frac{f(D^c|\theta_l, M_l)f(\theta_l|M_l)}{f(D^c|M_l)} \quad (4)$$

Below two sections present the calibration illustration for case of corrosion rate measurement and crushing strength ( $f_c$ ) estimation from rebound number values, respectively.

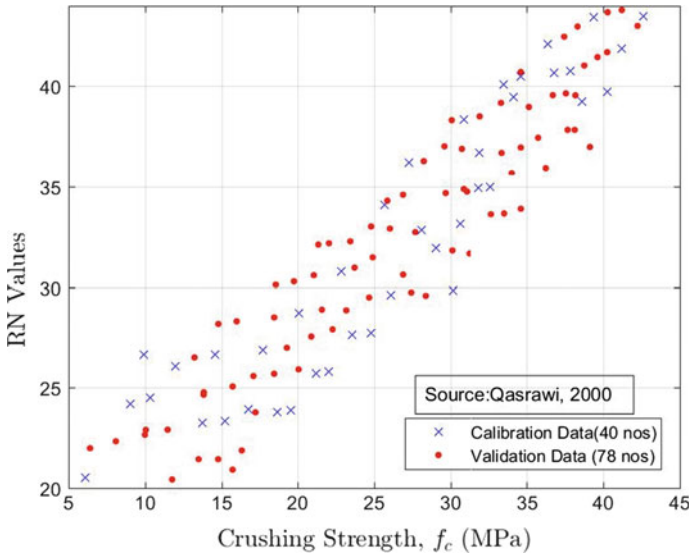
### 3.1 Rebound Hammer Calibration

Rebound hammer testing is an important step in structural auditing. The crushing strength ( $f_c$ ) is obtained from the graph given in the instrument manual or calculating from model available from the literature. The quality and accuracy of  $f_c$  estimation rely on the precision of the calibration technique and calibration data ( $D_n^c$ ). Literature studies show that a great amount of work is being carried out correlating rebound number (RN) and crushing strength ( $f_c$ ) of concrete [19, 20]. Linear law, bilinear law, power law, double power law, etc. are being used to relate RN and  $f_c$  [1]. Codes and standards also suggest calibration of rebound hammer instrument using suitable technique and also provide calibration curves [7, 12, 21]. In this section, Bayesian updating is demonstrated for calibration of rebound hammer instrument. In this study, linear law, power law, and exponential law are adopted as there exist more than 20 models calibrated by different researchers [19]. Equation for each model is given in Table 1.

In the above-given models,  $a$  and  $b$  are model parameters, and  $E$  is the random error discussed in above section. The training data is sourced from Qasrawi, 2000 [22], using the WebPlotDigitizer [8], 118 points set of pairs of the corresponding  $f_c$  and RN were extracted. These were bifurcated into “Training Data” ( $D_n^c$ ) and “Validation Data” ( $D_n^v$ ) as show in Fig. 1. Validation Data is used for model validation which is beyond the scope of this paper. The prior parameters ( $a, b, E$ ) have to be defined in such a way that  $f_c$  obtained should not be negative. For e.g., only in linear model,  $a$  can take a negative value. In a linear model, the parameters  $b$  and  $\sigma$  can take only positive values. To keep them positive, we rewrite them as:  $b = e^{b_1}$ ,  $\sigma = e^s$ . Thus, the parameters in case of linear law are [ $a b_1 s$ ]. Similarly for power and exponential model, the parameters are [ $a_1 b_1 s$ ]. The prior parameters in this study considered as non-informative and a standard normal distribution  $N(\mu = 0, \sigma = 1)$  are assumed [8]. The joint prior distribution function (JPDF) is taken as the product of the densities of individual parameters, by assuming each parameter to be independent of others. Effect of prior in the presence of large number of training data is negligible; however, calibration results are influenced for small number of training data. Likelihood is formulated according to Eq. (2). Posterior was evaluated using Eq. (4) and samples set of parameters were obtained. In the MCMC simulation, a Gaussian proposal density is chosen:  $q(\Theta^+|\theta_m) \sim N(\theta_m, C)$ . Here, C is an appropriate diagonal covariance matrix, chosen to ensure sufficient mixing of the Markov Chain [23]. Covariance matrix has to be so tuned that the acceptance rate should lie between 0.20 and 0.35 for three parameter [24]. As it was difficult to tune the covariance matrix manually, DRAM was used to obtain C by running it for an initial run of  $N = 2000$  or

**Table 1** Models relating rebound number (R) and crushing strength of concrete

Linear Law	$f_c = a + b.RN + E$
Power Law	$f_c = a.RN^b + E$
Exponential Law	$f_c = a.exp(b.RN) + E$



**Fig. 1** Relationship of rebound number (R) and crushing strength of concrete ( $f_c$ ) [22]

$N = 5000$  simulations. The covariance matrix obtained from DRAM was then used in Metropolis–Hastings (MH) algorithm for generating samples of  $N = 10,000$  of each parameter [8]. Now to check diagrammatically the bounds and the fitting of the calibration (“training”) data, we carry out a forward problem for prediction of crushing strength  $f_c$  versus the same calibration data. Figures 2, 3 and 4 show the prediction of each model using the calibrated data with 95% bounds. To carry out this procedure, we substitute all  $N$  samples of for particular  $R_i$  measurement reading, thus we get  $N = 10,000$   $f_c$  values for one NDT reading. Mean is evaluated and two times the standard deviation is plotted for 95% bounds.

## 4 Bayesian Model Selection

As discussed in Sect. 3.1, three models were selected for Bayesian calibration, and thus there can be large number of varied types of model available. Models with less number of parameters are called “simple” while with more number of parameters is called “complex.” It is necessary to select a model which is a compromise between model complexity and goodness-of-fit, and fits best to the calibration data. The appropriateness model can be selected based on criteria such as: checking the difference of error between the measured value and the predicted value of the model or curve fitting. In this study, model selection is carried out using Bayes’ theorem.

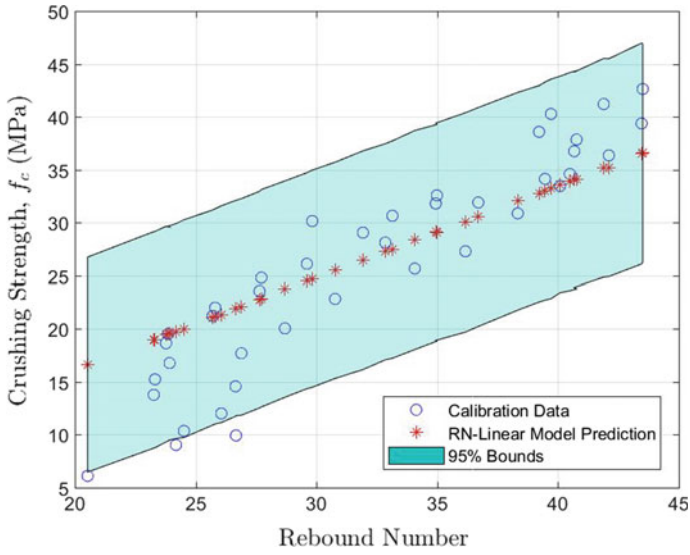


Fig. 2 Model Prediction of  $f_c$  versus the calibration data ( $D_n^c$ ) for linear model

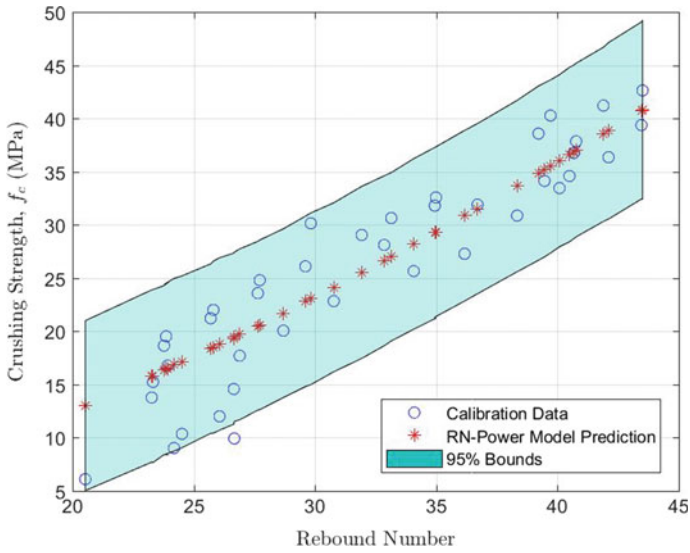
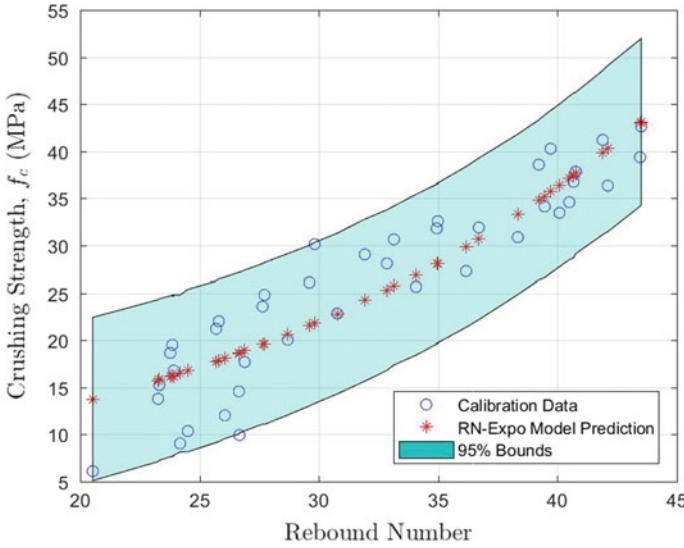


Fig. 3 Model Prediction of  $f_c$  versus the calibration data ( $D_n^c$ ) for power model



**Fig. 4** Model Prediction of  $f_c$  versus the calibration data ( $D_n^c$ ) for exponential Model

In Sect. 3.2, we have three a priori models ( $M_l$ ), the probability of a particular model  $M_l$ , conditioned on  $\mathbf{M}$  and the data  $D_n^c$ , can be obtained by Bayes’ theorem [14],

$$p(M_l|D_n^c, \mathbf{M}) = \frac{f(D_n^c|M_l) \cdot p(M_l|\mathbf{M})}{\sum_{l=1}^L f(D_n^c|M_l) \cdot p(M_l|\mathbf{M})} \tag{5}$$

where  $p(M_l|D_n^c, \mathbf{M})$  is the prior probability of each model  $M_l$ , which relates the plausibility of each model, such that  $\sum p(M_l|\mathbf{M}) = 1$ . The term  $f(D_n^c|M_l)$  is called as the evidence, and its calculation is generalized by Cheung and Beck, 2010 [25]. The best model maximizes the posterior model plausibility  $p(M_l|D_n^c, \mathbf{M})$ , compared to other models. This can be extended that the numerator of Eq. (5) should be maximized. The number of models available for a particular law was random in number and taking into consideration that these laws can be infinitely fitted, the prior probability is considered to be equal:  $p(M_l|\mathbf{M}) = 1/3$ . This can also be calculated based on the ratio of total number of each model available, but this is very uncertain in this case hence we take it equal.

Application of Bayes’ theorem enforces the belief that it is better to avoid undesired extra effort which can be done in few steps. This belief is carried out when the evidence is written as [14],

$$\ln f(D_n^c|M_l) = E [\ln f(D_n^c|\theta, M_l)] - E \left[ \ln \frac{f(\theta|D_n^c, M_l)}{f(\theta|M_l)} \right] \tag{6}$$



**Table 2** Model selection results

Models from Sect. 3.2	Prior probability	Rank-based prior probability	Log evidence	Expected information gain
$M_l$	$p(M_l D_n^c, \mathbf{M})$		$\ln f(D_n^c M_l)$	$E \left[ \ln \frac{f(\theta D_n^c, M_l)}{f(\theta M_l)} \right]$
Linear Law	0.03681	2	-128.48	6.81
Power Law	0.9675	1	-122.92	11.02
Exponential Law	0.02881	3	-126.43	12.29

where  $E[\cdot]$  is the expectation with respect to the posterior  $f(\theta|D_n^c, M_l)$ . The first term on the right-hand side of Eq. (6) is a measure of the data-fit and the second term represents the Kullback–Leibler information gain from the data. The information gain is higher for complex models as they over-fit the data, thereby penalizing the evidence of a complex model [8]. Thus, the evidence is a trade-off between how well the model fits the data versus its complexity. Based on the posterior model probability Eq. (5), the models are ranked in Table 2 and provide the values of the log evidence and expected information gain.

Thus from Table 2, it can be concluded that power law strikes the best figure for the most plausible model and expected information gain with respect to other models. Thus using Bayesian model selection technique, comparing two parameters, the best model can be inferred.

## 5 Conclusions

NDT instrument calibration is a vital step for good quality assessment of the concrete properties. This study provides the framework of probabilistic approach adopting Bayesian model calibration which is superior to deterministic method as it takes into account the uncertainty involved. The key conclusions are summarized below:

1. Bayesian calibration results in sample values of parameters rather than single point value evaluated using deterministic analysis. Hence a PDF is obtained for the final estimate which gives better reliability on the prediction/estimation of concrete properties.
2. Calibration parameters can be continuously updated with new incoming data which will reduce the uncertainty involved in the measurement phenomenon.
3. Concrete properties of old RC structures having high spatial variation, probabilistic technique provides a technique to integrate the structural properties as a whole.
4. Bayesian model selection provides with an approach which evaluates the best model which is simple and fits the calibration data.

## 6 Future Scope

1. In this work, only three model laws were selected for calibration. More number of complex models can be incorporated in the calibration and model selection process. Also, bivariate models can be taken into consideration.
2. Model validation by challenging the calibrated model with a different set of data.
3. Data fusion of two different NDT instrument using Bayesian inference.

## References

1. V.M. Malhotra, N.J. Carino, *Handbook on Nondestructive Testing of Concrete*, 2nd edn. (CRC Press, Boca Raton, FL, 2003)
2. M. Shetty, *Concrete Technology* (Chand Co. LTD, 2005), pp. 420–453
3. S.K. Verma, S.S. Bhadauria, S. Akhtar, In-situ condition monitoring of reinforced concrete structures. *Front. Struct. Civ. Eng.* **10**(4), 420–437 (2016)
4. BS:1881-202, *Recommendations for surface hardness testing by rebound hammer*. BSI Publication, London (1986)
5. BS:1881-203, *Recommendations for measurement of ultrasonic pulse for concrete*. BSI Publication, London (1986)
6. BIS 13311, *Non-Destructive Testing of concrete—Methods of Test : Part 1—Ultrasonic Pulse Velocity*. Bureau of Indian Standards, New Delhi, India (1992)
7. BIS 13331, *Non-Destructive Testing of concrete—Methods of Test : Part 2—Rebound Hammer*. Bureau of Indian Standards, New Delhi, India (1992)
8. S.A. Faroz, *Assessment and Prognosis of Corroding Reinforced Concrete Structures through Bayesian Inference*. Ph.D. Thesis, Indian Institute of Technology Bombay, Mumbai, India (2017)
9. P.K. Mehta, Durability of concrete—fifty years of progress? *Spec. Publ.* **126**, 1–32 (1991)
10. M. Ploix, V. Garnier, D. Breyse, J. Moysan, NDE data fusion to improve the evaluation of concrete structures. *NDT E Int.* **44**(5), 442–448 (2011)
11. Z. Sbartai, D. Breyse, M. Larget, J. Balayssac, Combining NDT techniques for improved evaluation of concrete properties. *Cem. Concr. Compos.* **34**(6), 725–733 (2012)
12. B. EN:13791, *Assessment of Insitu Compressive Strength in Structures and Precast Concrete Components* (2007)
13. C. Lu, W. Jin, R. Liu, Reinforcement corrosion-induced cover cracking and its time prediction for reinforced concrete structures. *Corros. Sci.* **53**(4), 1337–1347 (2011)
14. M. Muto, J.L. Beck, Bayesian updating and model class selection for hysteretic structural models using stochastic simulation. *J. Vib. Control* **14**(1–2), 7–34 (2008)
15. J.L. Beck, K.-V. Yuen, Model selection using response measurements: Bayesian probabilistic approach. *J. Eng. Mech.* **130**(2), 192–203 (2004)
16. S.A. Faroz, N.N. Pujari, R. Rastogi, S. Ghosh, Risk analysis of structural engineering systems using Bayesian inference, in *Modeling and Simulation Techniques in Structural Engineering* (IGI Global, 2017), pp. 390–424
17. S.A. Faroz, N.N. Pujari, S. Ghosh, A Bayesian Markov Chain Monte Carlo approach for the estimation of corrosion in reinforced concrete structures, in *Proceedings of the Twelfth International Conference on Computational Structures Technology*, vol. 10, Stirlingshire, UK (2014)
18. A.H.-S. Ang, W.H. Tang, *Probability Concepts in Engineering Planning and Design* (1984)
19. D. Breyse, Nondestructive evaluation of concrete strength: An historical review and a new perspective by combining NDT methods. *Constr. Build. Mater.* **33**, 139–163 (2012)

20. E. Arioğlu, O. Koyluoglu, Discussion of prediction of concrete strength by destructive and nondestructive methods by Ramyar and Kol. *Cement and Concrete World* (1996)
21. ASTM:C805/C805 M-13a, *Test for Rebound Number of Hardened Concrete*. ASTM International, West Conshohocken, PA (2013)
22. H.Y. Qasrawi, Concrete strength by combined nondestructive methods simply and reliably predicted. *Cem. Concr. Res.* **30**(5), 739–746 (2000)
23. S. Chib, E. Greenberg, W.R. Gilks, S. Richardson, D. Spiegelhalter, Markov chain Monte Carlo in practice. *Am. Stat.* **49**(4), 327–335 (1995)
24. G. Roberts, J.S. Rosenthal et al., Optimal scaling for various Metropolis-Hastings algorithms. *Stat. Sci.* **16**(4), 351–367 (2001)
25. S.H. Cheung, J.L. Beck, Calculation of posterior probabilities for Bayesian model class assessment and averaging from posterior samples based on dynamic system data. *Comput. Aided Civ. Infrastruct. Eng.* **25**(5), 304–321 (2010)