22nd International Conference
ENGINEERING MECHANICS 2016
Svratka, Czech Republic, 9 – 12 May 2016

EPISTEMIC UNCERTAINTY IDENTIFICATION VIA DIFFERENT BAYESIAN INFERENCE METHODS

E. Janouchová[*], A. Kučerová[**], J. Sýkora[***]

Abstract: *Currently the Bayesian inference becomes increasingly popular and more widespread approach to parameter identification, which allows to estimate values of input parameters together with appropriate uncertainties by combining a priori information and experimental measurements. In this case, observations are assumed to be performed for the specific yet unknown values of input parameters and epistemic uncertainty arising from experimental errors and lack of knowledge is reduced with an increasing number of experimental observations. Bayes' rule provides an elegant solution to this inverse problem by making it well-posed. However the resulting a posteriori probability distribution usually has a complicated formulation, which cannot be treated analytically. To overcome this obstacle, several methods were developed. The most commonly referred techniques in literature are based on the Markov chain Monte Carlo method, less mentioned approaches utilize the Kalman filter or optimal transport maps. The aim of this contribution is to review and compare these methods of the Bayesian inference.*

Keywords: Epistemic uncertainty, Bayesian inference, Markov chain Monte Carlo, Kalman filter, Optimal transport map.

1. Introduction

In order to predict the behaviour of the structural system under the loading in a computational way, the corresponding numerical model has to be properly calibrated. In other words, parameters of the mathematical model of the system have to be estimated as accurately as possible to obtain realistic predictions, e.g. for usage in an appropriate reliability analysis or structural design optimisation. To infer the model parameters from indirect experimental measurements one can proceed in two principally different ways. The traditional approach is deterministic, while advances in surrogate modelling and increasing computational capacity of modern computers permitted many researches to focus on parameter identification in probabilistic setting.

The most common method of parameter estimation is based on fitting the response of the numerical model to the experimental data. This deterministic approach leads to optimising parameters so as to minimise the difference between the data and the model response. The optimisation problem is, however, often ill-posed and thus requires the employment of robust optimisation algorithms. The result of such an optimisation process is only the single-point estimate of parameter values, as you can see in Fig. 1b, thus any information beyond the mean values of parameters is omitted. Consequently, this deterministic inversion method does not provide any quantification of the uncertainty in parameter estimates which in fact exists and is caused by e.g. an insufficient number of observations and measurement errors.

In this context, uncertainties can be divided into two main categories according to whether a source of nondeterminism is irreducible or reducible (Oberkampf et al., 2002). This contribution focuses on identification of epistemic (reducible, subjective, cognitive) uncertainty arising from our lack of knowledge which is supposed to be reduced by any new measurement according to the coherence of

[*] Ing. Eliška Janouchová: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, eliska.janouchova@fsv.cvut.cz

[**] Ing. Anna Kučerová, PhD: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, anicka@cml.fsv.cvut.cz

[***] Ing. Jan Sýkora, PhD: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, jan.sykora.1@fsv.cvut.cz

learning (Mantovan & Todini, 2006; Beven et al., 2007). While the inherent stochasticity is expressed by aleatory uncertainty which cannot be reduced.

In the last decades probabilistic methods for stochastic modelling of uncertainties have become applicable thanks to a growing computational capacity of modern computers. The probabilistic approach restates the inverse problem as well-posed in an expanded stochastic space by modelling the parameters as well as the observations as random variables with their probability distributions (Kaipio & Somersalo, 2005). Several methods for the uncertainty quantification in probabilistic settings have been proposed in the literature. The last decade witnessed an intense development in the field of Bayesian updating of epistemic uncertainty (Fig. 1c) in description of deterministic material or structural parameters, see e.g. Marzouk et al. (2007). Here, a likelihood function is established to quantify our confidence in observed data, with the goal to update our prior knowledge on model parameters (Gelman et al., 2004). The increasing popularity of Bayesian methods is motivated by developments in the field of spectral stochastic finite element method, which allows to alleviate the computational burden by surrogate models such as polynomial chaos expansions (Marzouk & Najm, 2009). The most commonly referred techniques of Bayesian inference in literature are based on the Markov chain Monte Carlo method (Marzouk et al., 2007), less mentioned approaches utilize the Kalman filter (Rosić et al., 2013) or optimal transport maps (El Moselhy & Marzouk, 2012). The aim of this contribution is to review and compare these methods of the Bayesian inference.

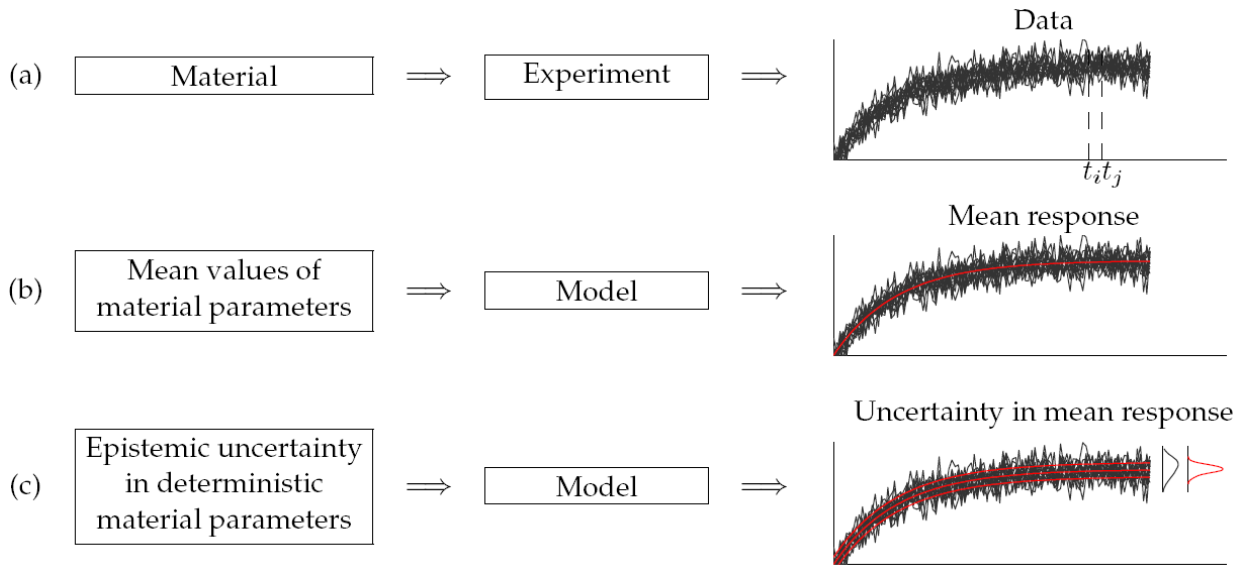

Fig. 1: Scheme of an experiment and different approaches to parameter identification.

2. Bayesian inference

Consider a stochastic problem

$$\boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\omega}) = \boldsymbol{y}(\boldsymbol{x}) + \varepsilon(\boldsymbol{\omega}) \tag{1}$$

with uncertain model parameters $\boldsymbol{x}$ and random observable data $\boldsymbol{z}$, which can be predicted by a model response $\boldsymbol{y}(\boldsymbol{x})$ besides a measurement error ε. In Bayesian statistics, probability represents a degree of belief about the parameter values (Tarantola, 2005). Combining the initial knowledge in the form of the prior distribution $p(\boldsymbol{x})$ and the experimental data as the likelihood function $p(\boldsymbol{z} \mid \boldsymbol{x})$ according to Bayes' rule

$$p(\boldsymbol{x}|\boldsymbol{z}) = \frac{p(\boldsymbol{z}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{z})} = \frac{p(\boldsymbol{z}|\boldsymbol{x})p(\boldsymbol{x})}{\int p(\boldsymbol{z}|\boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x}}, \tag{2}$$

we obtain the posterior distribution of the parameters. The mean values of the updated distribution are equal to the best guess of the parameters' values with the uncertainty represented by the corresponding variance. However the posterior statistical moments cannot be generally computed analytically, because the identified distribution including the whole structural model is too complicated. To overcome this obstacle, several methods were developed.

2.1. Markov chain Monte Carlo

Markov chain Monte Carlo (MCMC) is a sampling method based on a creation of an ergodic Markov chain of required stationary distribution equal to the posterior (Gilks, 2005; Geyer, 2011). There are different algorithms for constructing this chain (Spall, 2003), e.g. Gibbs sampler or Metropolis-Hastings algorithm, which avoids calculating of the normalisation constant in Eq. (2) by evaluating only ratios of target probabilities. Suitable setting of the proposal distribution for a random walk is important and can be evaluated on the basis of acceptance rate (Rosenthal, 2011) or autocorrelation which is required to be minimal. The convergence speed of the procedure depends also on the appropriate choice of the starting point (Geyer, 2011). The essential advantage of this method is its versatility for usage with nonlinear models, when for an infinite number of samples it gives the exact solution. The disadvantage of this method is its high computational effort resulting from necessity of a high number of model simulations. In order to accelerate this sampling procedure in identification process, the evaluations of a numerical model can be replaced by evaluations of a computationally efficient model surrogate.

2.2. Kalman filter

The second way of obtaining updated posterior distribution comprises Bayesian linear methods, see e.g. Rosić et al. (2016), based on Kalman filtering (Kálmán, 1960). The basic idea of these methods is to update the prior random variable X_f by a linear map to a linear Bayesian posterior estimate

$$X_a = X_f + K\left(\boldsymbol{z}(\boldsymbol{x},\boldsymbol{\omega}) - Y_f\left(X_f\right)\right), \tag{3}$$

where Y_f is the prior model response and the Kalman gain

$$K = \mathbf{C}_{X_f Y_f}\left(\mathbf{C}_{Y_f} + \mathbf{C}_{\varepsilon}\right)^{-1} \tag{4}$$

is computed from the corresponding covariance matrices and measurement covariance $\mathbf{C}_{\varepsilon}$.

The posterior X_a can be estimated by so called the ensemble Kalman filter algorithm based on updating of prior Monte Carlo samples, which also serve for computation of the covariance matrices. The method requires a smaller number of samples than previous MCMC method, but the identification of the uncertainty is not generally so accurate as with MCMC. Another approach is to approximate the random variables by polynomial chaos expansions, which enables to evaluate the Kalman gain and posterior X_a in an algebraic way (Rosić et al., 2012). Its main advantage is elimination of computationally demanding model simulations. However, the result is exact only in a special case of a linear model and normally distributed random variables, in another cases these methods are only approximate.

2.3. Optimal maps

This technique is based on formulation of a transport function or a map which transforms the prior random variable X_f into the posterior random variable X_a (El Moselhy & Marzouk, 2012) and arises from the context of optimal transport theory. The authors describe the map by multivariate orthogonal polynomials and the solution is obtained by the optimisation of the corresponding polynomial coefficients. The cost function is defined with a help of the Kullback-Leibler divergence expressing the discrepancy between the prior density $p(\boldsymbol{x})$ and approximate map-dependent prior density. The prior has to be expressible by standard random variables whose probability distribution is orthogonal to the chosen polynomial basis (Xiu & Karniadakis, 2002). The posterior is then identified in the form of polynomial chaos expansion which is efficient in terms of analytical evaluation of posterior statistical moments. Thanks to deterministic expression of the map, one can easily sample from the posterior by transforming the prior samples.

3. Conclusions

The contribution is focused on presenting different numerical methods of Bayesian inference, which provides estimation of unknown model parameters along with a probability description of epistemic uncertainty corresponding to noisy experimental observations.

The most universal and well-known method is MCMC, which however, suffers from high computational demands. The sampling procedure can be accelerated by using a surrogate model instead

of the full numerical model, but it brings a possible source of approximation error. The ensemble Kalman filter needs less model simulations than MCMC, but the identified uncertainty does not reach so great precision because of the problem linearisation. The polynomial chaos based variant of linear Bayesian update avoids time-consuming sampling, but its accuracy also depends on the nonlinearity of the investigated problem. The last considered method of probabilistic parameter estimation is based on optimising deterministic mapping between prior and posterior. The approach is suitable for nonlinear and high-dimensional problems, the essential part of the procedure is a creation of the map composition and optimisation of its parameters. Formulation in terms of orthogonal polynomials allows to compute posterior statistical moments and to generate posterior samples very efficiently.

Acknowledgement

The financial support of the Ministry of Education, Youth and Sports of the Czech Republic (Project No. 8F15004), the Czech Science Foundation (Project No. 16-11473Y) and the Grant Agency of the Czech Technical University in Prague (Project No. SGS16/037/OHK1/1T/11) is gratefully acknowledged.

References

Beven, K., Smith, P. & Freer, J. (2007) Comment on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology" by Pietro Mantovan and Ezio Todini. Journal of Hydrology, 338, pp. 315-318.

El Moselhy, T. A. & Marzouk, Y. M. (2012) Bayesian inference with optimal maps. Journal of Computational Physics, 231, 23, pp. 7815-7850.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (2004) Bayesian data analysis. Chapman & Hall/CRC.

Geyer, C. J. (2011) Handbook of Markov Chain Monte Carlo. Chapman & Hall/CRC.

Gilks, W. R., Deland, S. M., Rutherford, B. M., Diegert, K. V. & Alvin, K. F. (2005) Markov Chain Monte Carlo. Encyclopedia of Biostatistics, Chichester, UK: John Wiley, 75(3), pp. 333-357.

Kaipio, J. & Somersalo, E. (2005) Statistical and Computational Inverse Problems. New York: Springer-Verlag, Applied mathematical sciences, 339 p.

Kálmán, R. E. (1960) A new approach to linear filtering and prediction problems. Transactions of the ASME-J. of Basic Engineering (Series D), 82, pp. 35-45.

Mantovan, P. & Todini, E. (2006) Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. Journal of Hydrology, 330, pp. 368-381.

Marzouk, Y. M. & Najm, H. N. (2009) Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. Journal of Computational Physics, 228, pp. 1862-1902.

Marzouk, Y. M., Najm, H. N. & Rahn, L. A. (2007) Stochastic spectral methods for efficient Bayesian solution of inverse problems. Journal of Computational Physics, 224, 2, pp. 560-586.

Oberkampf, W. L., DeLand, S. M., Rutherford, B. M., Diegert, K. V. & Alvin, K. F. (2002) Error and uncertainty in modeling and simulation. Reliability Engineering and System Safety, 75, pp. 333-357.

Rosenthal, J. S. (2011) Optimal Proposal Distributions and Adaptive MCMC, in: Handbook of Markov Chain Monte Carlo. Chapman & Hall/CRC, pp. 93-140.

Rosić, B., Kučerová A., Sýkora J., Pajonk O., Litvinenko, A & Matthies, H. G. (2013) Parameter Identification in a Probabilistic Setting. Engineering Structures, 50, pp. 179-196.

Rosić, B. V., Litvinenko, A., Pajonk, O. & Matthies, H. G. (2012) Sampling-free linear Bayesian update of polynomial chaos representations. Journal of Computational Physics, 231, pp. 5761–5787.

Rosić, B., Sýkora J., Pajonk O., Kučerová A. & Matthies, H. G. (2016) Comparison of Numerical Approaches to Bayesian Updating in Computational Methods for Solids and Fluids, 41, pp.427-461.

Spall, J. C. (2003) Estimation via Markov Chain Monte Carlo. IEEE Control Systems Magazine, 23, 2, pp. 34-45.

Tarantola, A. (2005) Inverse problem theory and methods for model parameter estimation. Philadelphia, PA: Society for Industrial and Applied Mathematics, 348 p.

Xiu, D. & Karniadakis, G. E. (2002) The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations. SIAM Journal on Scientific Computing, 24, 2, pp. 619-644.