Probabilistic Inversion for Estimating the Variability of Material Properties

A Bayesian Multilevel Approach

Author(s):
Nagel, Joseph; Sudret, Bruno

Publication date:
2013

Permanent link:
https://doi.org/10.3929/ethz-a-010034843

Originally published in:
Proceedings of the 11th International Probabilistic Workshop, Brno 2013

Probabilistic Inversion for Estimating the Variability of Material Properties: A Bayesian Multilevel Approach

Joseph Nagel [a], Bruno Sudret [a]

[a]Chair of Risk, Safety & Uncertainty Quantification,
Institute of Structural Engineering,
ETH Zürich, Switzerland
nagel@ibk.baug.ethz.ch, sudret@ibk.baug.ethz.ch

Abstract: Bayesian inference provides a convenient framework for the solution of inverse problems and the purposes of uncertainty quantification. It allows the estimation of model parameters from noisy data in a large variety of experimental situations. A major asset of the Bayesian paradigm is its unique and consistent treatment of complex models. Within the realms of Bayesian multilevel models aleatory variability and epistemic uncertainty can be naturally accounted for. In this contribution we will present the joys and sorrows of Bayesian multilevel modeling with a focus on probabilistic inversion, i.e. the inference of the variability of model inputs throughout a number of similar experiments. The possibilities it opens up as well as the problems it suffers from will be presented in the context of engineering applications. To that end we devise a simple academic example within the domain of structural engineering. On the basis of this simple problem we will develop the Bayesian multilevel perspective and exemplify its computational machinery. Under consideration is an ensemble of simply supported beams which are subjected to concentrated point loads. Measuring deflections allows for the identification of the variability of Young's moduli across the sample of beams.

Keywords: inverse problems, Bayesian inference, multilevel/hierarchical modeling, probabilistic inversion, Markov chain Monte Carlo, data augmentation, simply supported beams

1 Inverse Problems

Broadly speaking solving an *inverse problem* is the inference of model parameters or the identification of system states from noisy measurements. This covers a wide range of problems that arise in nearly all branches of natural science and engineering. Driven by the rise of powerful general-purpose computing facilities and the availability of dedicated software environments the Bayesian approach has become prevalent. It establishes a rigorous probabilistic framework and ships with a vast set of dedicated computational techniques. Scarce data and available expert information can be easily dealt with in Bayesian data analysis. A manifest advantage of the Bayesian frame is that it inherently is a means to *uncertainty quantification*. The result of any full Bayesian parameter estimation is basically a measure of epistemic (un)certainty over a set of admissible values.

1.1 Parameter Estimation

The generic *parameter estimation* problem can be stated as follows. A physical forward model $\mathcal{M}$ describes the system or phenomena under consideration and a fixed yet unknown model parameter $\boldsymbol{x}$ is to be identified in an experiment. Such a model may be viewed as a map $\boldsymbol{x} \in \mathcal{D}_{\boldsymbol{x}} \subset \mathbb{R}^p \mapsto \boldsymbol{y} = \mathcal{M}(\boldsymbol{x}) \in \mathbb{R}^q$. In engineering applications it can be a merely analytical expression, the numerical solution of a set of partial differential equations or even a surrogate model [8]. Subjected to known experimental conditions $\boldsymbol{d}_i$ a set of observations $\boldsymbol{y}_i$ is acquired for $i = 1, \dots, n$. Thus for parameter estimation the prototype of a statistical data model is written as

$$\boldsymbol{y}_i = \mathcal{M}(\boldsymbol{x}, \boldsymbol{d}_i) + \boldsymbol{\varepsilon}_i. \tag{1}$$

The residual $\boldsymbol{\varepsilon}_i \sim f_{\boldsymbol{E}}(\boldsymbol{\varepsilon}_i)$ accounts for the discrepancy between data $\boldsymbol{y}_i$ and model predictions $\mathcal{M}(\boldsymbol{x}, \boldsymbol{d}_i)$ due to measurement errors, numerical approximations as well as model inadequacies. The most widespread choice for a residual distribution is Gaussian one $f_{\boldsymbol{E}}(\boldsymbol{\varepsilon}_i) = \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$. The likelihood function is given as

$$\mathcal{L}(\boldsymbol{y}_1, \dots, \boldsymbol{y}_n | \boldsymbol{x}) = \prod_{i=1}^{n} f_{\boldsymbol{E}}\big(\boldsymbol{y}_i - \mathcal{M}(\boldsymbol{x}, \boldsymbol{d}_i)\big). \tag{2}$$

The Bayesian approach to parameter estimation consists in updating a prior or expert knowledge $\pi(\boldsymbol{x})$ in the light of the data $\boldsymbol{y}_i$. The elicitation of this prior distribution is a delicate issue in any Bayesian data analysis. By conditioning on the data the result of Bayesian inference is the posterior density

$$\pi(\boldsymbol{x} | \boldsymbol{y}_1, \dots, \boldsymbol{y}_n) \propto \mathcal{L}(\boldsymbol{y}_1, \dots, \boldsymbol{y}_n | \boldsymbol{x})\, \pi(\boldsymbol{x}). \tag{3}$$

This density quantifies an epistemic a posteriori degree of plausibility. Since it is analytically-closed only on rare occasions one commonly relies on Monte Carlo (MC) sampling from it [7]. Markov chain Monte Carlo (MCMC) can provide powerful mechanisms to draw samples from the posterior by constructing a suitable Markov chain over its support. This only requires the evaluation of posterior density at a number of specific points.

1.2 Probabilistic Inversion

Beyond Bayesian updating of an unknown model parameter another interesting class of problems is known as *probabilistic inversion*. Instead of attributing the discrepancy in the data $\boldsymbol{y}_i = \mathcal{M}(\boldsymbol{x}_i, \boldsymbol{d}_i) + \boldsymbol{\varepsilon}_i$ solely to a residual $\boldsymbol{\varepsilon}_i$, one assumes an aleatory variability

$$(\boldsymbol{x}_i|\boldsymbol{\theta}) \sim f_{\boldsymbol{X}|\boldsymbol{\Theta}}(\boldsymbol{x}_i|\boldsymbol{\theta}) \tag{4}$$

over the set of experiments for $i = 1, \ldots, n$. The vector of hyperparameters $\boldsymbol{\theta}$ determines the variability of the model inputs $\boldsymbol{x}_i$. Instead of a fixed yet unknown model parameter $\boldsymbol{x}$ inferential interest concentrates on the unknown hyperparameters $\boldsymbol{\theta}$. One therefore constructs a marginalized likelihood

$$\mathcal{L}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n|\boldsymbol{\theta}) = \prod_{i=1}^{n} \int_{\mathcal{D}_{\boldsymbol{x}_i}} f_{\boldsymbol{E}}\big(\boldsymbol{y}_i - \mathcal{M}(\boldsymbol{x}_i, \boldsymbol{d}_i)\big)\, f_{\boldsymbol{X}|\boldsymbol{\Theta}}(\boldsymbol{x}_i|\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{x}_i. \tag{5}$$

by integrating out the intermediate variables $\boldsymbol{x}_i$. Depending on the context these quantities are also called latent variables or missing data. Having elicited prior information $\pi(\boldsymbol{\theta})$ the updated posterior follows as

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) \propto \mathcal{L}(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n|\boldsymbol{\theta})\, \pi(\boldsymbol{\theta}). \tag{6}$$

However, the evaluation of the marginalized likelihood Eq. (5) when sampling from the posterior Eq. (6) remains an issue. Indeed the numerical integration involves a large number of calls to the forward model $\mathcal{M}$ which may be computationally expensive to run.

2 Multilevel Modeling

As there seems to be no universal definition, we define here *multilevel* or *hierarchical models* as being "an assembly of submodels at different levels of a hierarchy", where conditional dependencies and deterministic maps between the quantities involved constitute the hierarchy. Indeed the problem outlined as probabilistic inversion is a multilevel problem. Treating it as such provides some conceptional insight and suggests computational enhancements. To summarize the overall model with a three-level hierarchy we write

$$(\boldsymbol{y}_i|\boldsymbol{x}_i) \sim f_{\boldsymbol{E}}\big(\boldsymbol{y}_i - \mathcal{M}(\boldsymbol{x}_i, \boldsymbol{d}_i)\big), \tag{7a}$$

$$(\boldsymbol{x}_i|\boldsymbol{\theta}) \sim f_{\boldsymbol{X}|\boldsymbol{\Theta}}(\boldsymbol{x}_i|\boldsymbol{\theta}), \tag{7b}$$

$$\boldsymbol{\theta} \sim \pi(\boldsymbol{\theta}). \tag{7c}$$

A directed acyclic graph (DAG) as in Fig. 1 provides an intuitive representation of this hierarchical model. Such graphs are widely known as Bayesian networks, too. Bayesian multilevel modeling provides the proper Bayesian framework for formulating complex inverse problems involving parameter variability and uncertainty together with a wealth of well-established computational techniques to solve them. It allows for a unique and consistent way of inference.

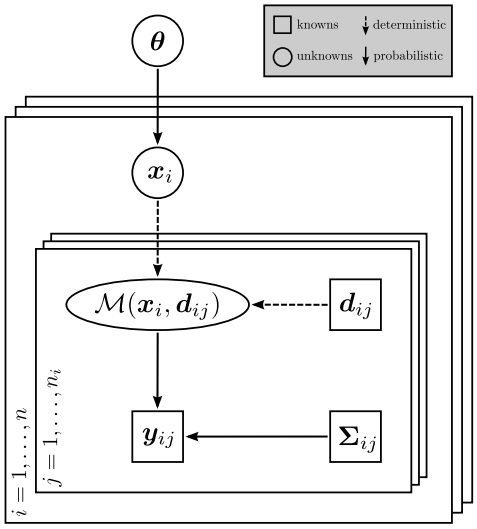


Figure 1: DAG of the generic framework. The overall multilevel model is assembled from probabilistic (⟶) and deterministic (⇢) relationships between unknown (○) and known (□) variables. The three main levels of variables are shown: the observable data $\boldsymbol{y}_{ij}$, the unobservable parameters $\boldsymbol{x}_i$ and the hyperparameter $\boldsymbol{\theta}$. Moreover there are known forward model inputs $\boldsymbol{d}_{ij}$ as well as known residual covariances $\boldsymbol{\Sigma}_{ij}$. There are $j = 1, \ldots, n_i$ observations for each individual unit $i = 1, \ldots, n$.

2.1 Types of Inference

Bayesian inference in hierarchical models is based on constructing the joint posterior $\pi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{\theta} | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ of all unknowns $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{\theta})$ conditioned on all knowns $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$. This posterior is given as

$$\pi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{\theta} | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) \propto \left(\prod_{i=1}^{n} f_{\boldsymbol{E}}\big(\boldsymbol{y}_i - \mathcal{M}(\boldsymbol{x}_i, \boldsymbol{d}_i)\big) \right) \left(\prod_{i=1}^{n} f_{\boldsymbol{X}|\boldsymbol{\Theta}}(\boldsymbol{x}_i | \boldsymbol{\theta}) \right) \pi(\boldsymbol{\theta}). \quad (8)$$

What is considered *nuisance* is then subsequently integrated out from the posterior. This is the unique Bayesian perspective to inference in multilevel models. Individual parameters $\boldsymbol{x}_i$ and/or the hyperparameter $\boldsymbol{\theta}$ can be of inferential interest.

On the one hand inference of individual $\boldsymbol{x}_i$ becomes possible by treating and estimating them as latent variables instead of integrating them out as nuisance. This allows for a joint learning mechanism that bases the estimation of $\boldsymbol{x}_i$ not only on the inversion of $\boldsymbol{y}_i$ but indirectly also on the information provided by $\boldsymbol{y}_j$ with $j \neq i$. In this context this is generally referred to as the optimal combination of information. The result is the posterior $\pi(\boldsymbol{x}_i | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ where the hyperparameter and all parameters but $\boldsymbol{x}_i$ have been marginalized out.

On the other hand we will concentrate here on the optimal and efficient estimation of the unknown hyperparameter $\boldsymbol{\theta}$. The correspondingly marginalized posterior density is given as

$$\pi(\boldsymbol{\theta} | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) = \int_{\mathcal{D}_{\boldsymbol{x}_1}} \ldots \int_{\mathcal{D}_{\boldsymbol{x}_n}} \pi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{\theta} | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) \, \mathrm{d}\boldsymbol{x}_1 \ldots \mathrm{d}\boldsymbol{x}_n. \quad (9)$$

It can be seen that constructing the joint posterior Eq. (8) and thereafter integrating out the intermediate parameters Eq. (9) is equivalent to Eq. (6) where they have already been integrated out in the likelihood Eq. (5).

2.2 MCMC Sampling

As far as the inference on $\boldsymbol{\theta}$ is concerned there is no conceptional difference between Eq. (6) and Eq. (9). However, the two schemes pose different numerical tasks and differ in their computational efficiency when it comes to MCMC sampling. While the former provides a relatively crude sampling scheme that involves the intricate marginal version of the likelihood, the latter suggests a sampling scheme that is based on the conditional densities $f_{\boldsymbol{E}}\big(\boldsymbol{y}_i - \mathcal{M}(\boldsymbol{x}_i, \boldsymbol{d}_i)\big)$ and $f_{\boldsymbol{X}|\boldsymbol{\Theta}}(\boldsymbol{x}_i|\boldsymbol{\theta})$. Constructing a Markov chain over the joint posterior Eq. (8) and discarding the samples of the intermediate variables $\boldsymbol{x}_i$ directly provides samples from the marginal posterior of $\boldsymbol{\theta}$. This way drawing but neglecting $\boldsymbol{x}_i$ facilitates to draw $\boldsymbol{\theta}$ from the marginal. While when normally stated as the output of a Bayesian analysis, marginals tend to hide dependencies between parameters, in our case this marginalization is indeed the final objective. That technique is sometimes called *data augmentation* [6] and has been trivially derived here for multilevel models. The key challenge posed lies now in the efficient sampling in a possibly high-dimensional parameter space.

2.3 Literature Review

The outlined framework represents some kind of an all-embracing superstructure that embeds formerly proposed approaches to handle variability and uncertainty in inverse problems as being particular facets or approximations of it. As a generic framework it allows to interrelate various established ideas and methods and it suggests some fundamentally new ones. Previously established approaches to probabilistic inversion subsume frequentist methods based on the explicit marginalization of the likelihood [4] and the approximate treatment based on linearization [2] or kriging [1] and expectation-maximization-type (EM) algorithms. Moreover there are approximate two-stage approaches that consist in the separate estimation of individual parameters and a subsequent direct statistical estimation of the hyperparameter [5] and Bayesian inference of the hyperparameter by means of MCMC and data augmentation [6]. A nice review of many methodologies for such and associated problems is found in [3].

3 Application Example: Simply Supported Beam

The system under consideration is a set of simply supported beams $i = 1, \ldots, n$ that are subjected to a pinpoint load F_i at midspan. The Young's modulus E_i is constant along the beam span of individual beams i. Individual beams i are assumed to be made out of the same material that exhibits an aleatory variability in E_i. Across the sample Young's moduli E_i are distributed according to a lognormal distribution $\mathcal{LN}(\lambda, \zeta)$ with mean value $\mu_E = \exp(\lambda + \zeta^2/2)$ and standard deviation $\sigma_E = \sqrt{\exp(2\lambda + \zeta^2)(\exp(\zeta^2) - 1)}$. We will assume here that the beam lengths L_i, widths b_i, heights h_i and the applied forces F_i are sufficiently well-known. These model inputs are therefore assigned deterministic values. For each beam the displacement $v_i(s_j)$ at positions $0 \leq s_j \leq L_i/2$ with $j = 1, \ldots, n_i$ along the

beam axis is given as

$$v_i(s_j) = \frac{F_i s_j}{48 E_i I_i} \left(3L_i^2 - 4s_j^2\right) . \tag{10}$$

For $L_i/2 \leq s_j \leq L_i$ a symmetric expression holds. Since the moment of inertia is given by $I_i = b_i h_i^3/12$, the maximal displacement at midspan $s_j = L/2$ is $v_i^{(\max)} = F_i L_i^3/48 E_i I_i$. Eq. (10) will constitute the forward model $v_i(s_j) = \mathcal{M}(E_i, F_i, L_i, I_i, s_j)$ for individual beams i. An individual beam is shown in Fig. 2, a simplified DAG for the full hierarchical problem is shown in Fig. 3. We envisage the experimental identification of how Young's moduli E_i are distributed across a sample of e.g. $n = 100$ timber beams. The procedure will be illustrated by conducting a numerical computer experiment. Therefore we fix and simulate the experimental setup to obtain synthetic pseudo-data and the available prior knowledge as follows.

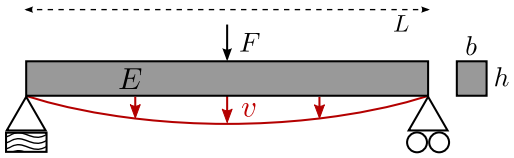


Figure 2: The simply supported beam

The beam dimensions are set to $L_i = 1\,\mathrm{m}$ and $b_i = h_i = 10\,\mathrm{cm}$ and the applied loads are chosen to be $F_i = 30\,\mathrm{kN}$. For $i = 1, \ldots, 100$ we sample E_i according to a lognormal distribution with $\mu_E = 15\,\mathrm{GPa}$ and $\sigma_E = 3\,\mathrm{GPa}$. This corresponds to a coefficient of variation $c = 20\,\%$. A set of pseudo-measurements for the displacements $v_i(s_j)$ at $n_i = 3$ arbitrarily chosen positions $s_1 = 25\,\mathrm{cm}$, $s_2 = 50\,\mathrm{cm}$ and $s_3 = 75\,\mathrm{cm}$ are generated for every i by perturbing the model predictions Eq. (10) with normally distributed and independent measurement noise $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2)$ with $\sigma_{ij} = 0.1\,\mathrm{mm}$. Component-wise the hyperprior on the hyperparameters $\boldsymbol{\theta} = (\theta_1, \theta_2)$ with a shape $\theta_1 = \mu_E$ and a scale parameter $\theta_2 = \sigma_E$ is set priorly independent $\pi(\boldsymbol{\theta}) = \pi(\mu_E)\,\pi(\sigma_E)$. The marginals are chosen to be the nearly non-informative uniform and proper prior distributions $\pi(\mu_E) \sim \mathcal{U}(0, 100)$ and $\pi(\sigma_E) \sim \mathcal{U}(0, 100)$.

4 Results

To the extent deemed reasonable we treat the forward model Eq. (10) as a representative black-box. For the sake of universality its role is to represent any other computational forward model, e.g. a FEM model, too. Approaches based on the analytical form of the model or possibly some linearization of it are therefore not considered. Given the described setup we will conduct the Bayesian probabilistic inversion by means of MCMC methods. For a start the probabilistic inversion is based on a simple random walk Metropolis sampling of the posterior Eq. (6). The results will be subsequently compared to a blockwise MCMC sampling scheme of the joint posterior Eq. (8). We rely on Matlab as the platform of our choice. A vectorized function implements the forward model Eq. (10) under consideration.

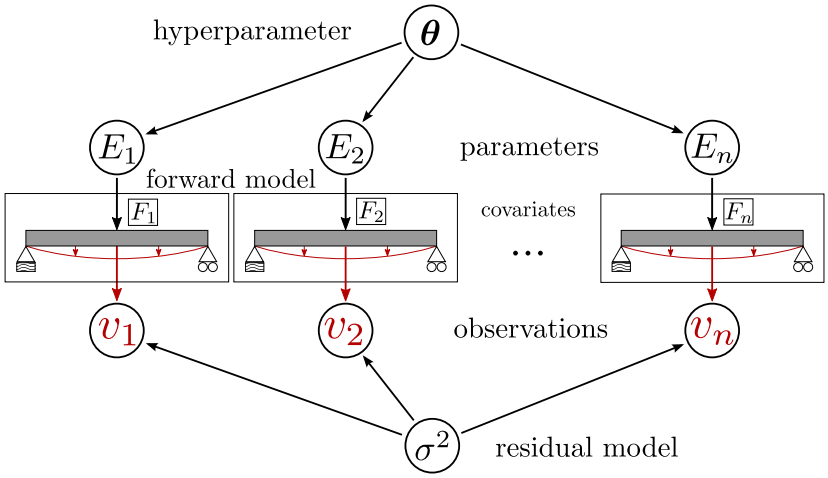


Figure 3: Simplified DAG of the simply supported beam example. The Young's modulus E_i is varying over a set of simply supported beams $i = 1, \ldots, n$ and a number of similar experiments is carried out to infer individual E_i or their variation $\boldsymbol{\theta} = (\mu_E, \sigma_E)$. The beam dimensions are assumed to be well-known and constant across the sample while the forces F_i act as covariates. A refinement of the residual model σ^2 allows to take more specific models σ_{ij}^2 into account.

4.1 Marginalized Likelihood

First of all we show the results of Bayesian probabilistic inversion that is based on the posterior Eq. (6). The integral in Eq. (5) is evaluated by means of simple Monte Carlo techniques, where the number of samples to simulate the likelihood is carefully set to the smallest yet acceptable number. This is accomplished by a systematic comparison to the results of an accurate numerical quadrature and MC sampling with an unreasonably high number of utilized samples. A sample size of $N = 10^5$ per factor in Eq. (5) was employed. The algorithm is tuned by reasonably setting the stepsize and initialized with a two-stage estimate $\hat{\boldsymbol{\theta}}$ of the hyperparameter. With a standard deviation of the Gaussian proposal density of $0.2\,\mathrm{GPa}$ an acceptance rate of ca. $65\,\%$ has been obtained. Traceplots of the resulting MCMC sample of μ_E and σ_E are shown in Fig. 4. The total sample size $N_{ML} = 10^3$ is chosen in such a way that the total execution time $t_{ML} = 1363\,\mathrm{s}$ is of similar magnitude as in the following run. Although being of crucial importance at this place we neither discuss any issues regarding convergence nor its diagnostics.

For the analysis a slightly optimized forward model Eq. (10) has been implemented. We take up the position that when comparing two methods their respective optimum should be compared together. One could also consider the employment of identically implemented forward models for different purposes a fairer comparison. For the sake of completeness we therefore run the same analysis again with the identical implementation of the forward model Eq. (10) which is not optimized for the integration Eq. (5) in particular. In this case the total execution amounted to ca. $t_{ML} = 1797\,\mathrm{s}$.

As the total runtime crucially depends on the number of utilized samples to do the MC integration we shortly state the execution times $t_{10^4} = 206\,\mathrm{s}$ and $t_{10^3} = 41\,\mathrm{s}$ for when $N_{ML} = 10^4$ or $N_{ML} = 10^3$ samples are employed, respectively. As in these cases the evaluation of

Eq. (5) turns out to be only insufficiently accurate, these numbers are just intended to provide a rough measure for if it were sufficient.

4.2 Joint Posterior Sampling

We carry out the analysis based on the joint posterior Eq. (8) and compare the results to the crude scheme above. A Markov chain over the $n + 2 = 102$-dimensional parameter space is constructed by means of a blockwise random walk Metropolis algorithm. The set of parameters $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ and the hyperarameter $\boldsymbol{\theta}$ constitute the different blocks. Tuning of the algorithm takes place by conveniently setting specific stepsizes for the different blocks. We use a Gaussian proposal density with standard deviations $0.02\,\mathrm{GPa}$ for the first block and $0.1\,\mathrm{GPa}$ for the second. This provides an acceptance rate of ca. $65\,\%$. Initialization is done with separate maximum likelihood estimates $\hat{\boldsymbol{x}}_i$ and the two-stage estimate $\hat{\boldsymbol{\theta}}$.

Traceplots of the Markov chain over μ_E and σ_E are given in Fig. 5 while Fig. 6 displays the corresponding simulated marginal posterior densities. This is done by conveniently binning the sample and normalizing the resulting histogram. The sampled two-dimensional posterior of $\boldsymbol{\theta}$ is shown in Fig. 7. For the sake of convenience the binning is the two-dimensional analogue of the one shown in Fig. 6 above. To produce the sample of size $N_{JP} = 10^6$ the execution time adds up to $t_{JP} = 1613\,\mathrm{s}$.

By comparison of Fig. 4 and Fig. 5, the sample sizes of $N_{ML} = 10^3$ and $N_{JP} = 10^6$ and the durations $t_{ML} = 1363\,\mathrm{s}$ and $t_{JP} = 1613\,\mathrm{s}$ the joint approach is shown to be superior to the integrated likelihood one. Having comparable mixing properties the two approaches differ in their computational demand and execution times, though. In our case the difference basically amounts to three orders of magnitude. When involving the integrated likelihood the executions time crucially depends on the exact method and specifications chosen. Nevertheless it seems plausible to generally conclude that these schemes suffer from the numerical burden posed by the integral.

5 Conclusion & Outlook

Bayesian inversion in multilevel models for the treatment of variability and uncertainty has been addressed. Multilevel models provide a convenient framework for probabilistic inversion and other types of inference whereas classical approaches can be considered partial or approximate treatments thereof. Here we have demonstrated the computational benefits of data augmentation for the purpose of probabilistic inversion. The variation of model inputs over a set of similar experiments has been identified in a multilevel analysis. The key challenge was the sampling of the marginal posterior of the hyperparameter that either involves an integrated likelihood or a high-dimensional parameter space. For that purpose the whole range of Bayesian computations as well as dedicated data augmentation sampling schemes are readily available. Future research plans involve the implementation of more efficient dedicated sampling techniques. Moreover the framework will be expanded by stochastic model inputs whose realizations are only imperfectly known but follow a known and prescribed

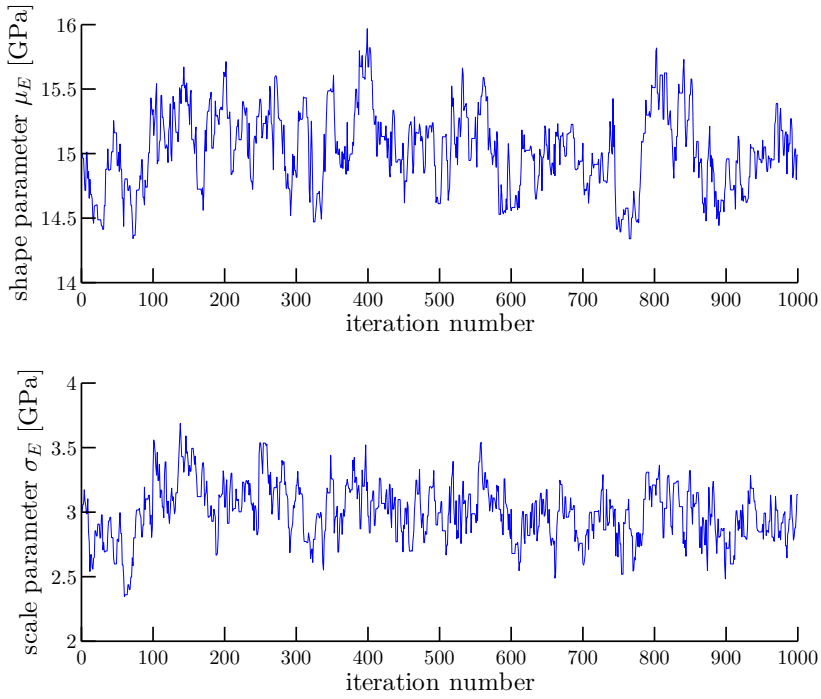


Figure 4: Traceplots of the Markov chain over μ_E and σ_E are shown for the crude probabilistic inversion. The total algorithm runtime was measured to be ca. $t_{ML} = 1363\,\mathrm{s}$. Mixing properties are comparable to the ones of the joint posterior approach below.

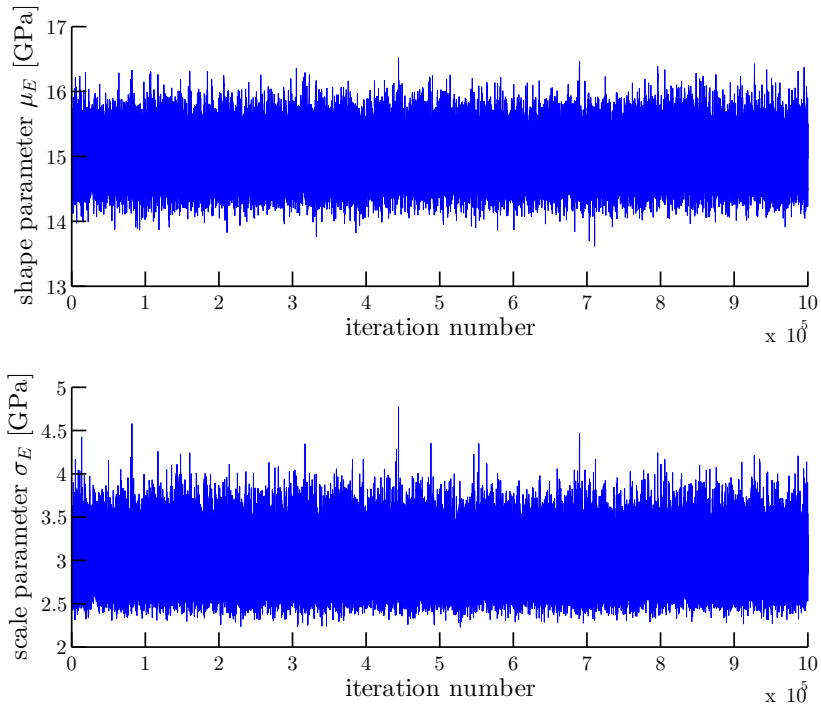
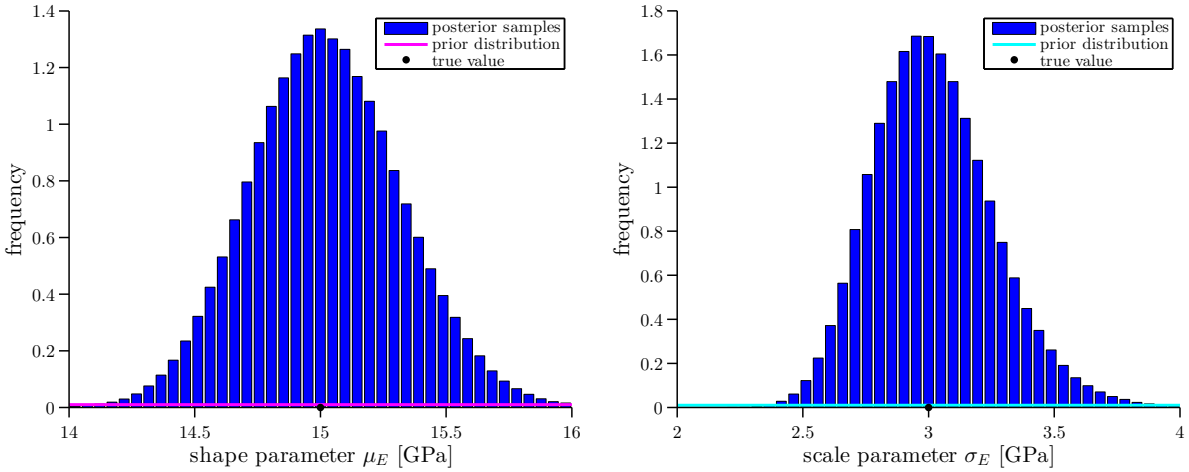


Figure 5: Traceplots of the Markov chain over the joint parameter space are depicted for μ_E and σ_E. Execution time has amounted to ca. $t_{JP} = 1613\,\mathrm{s}$. Since for comparable runtimes a much higher number of samples $N_{JP} \gg N_{ML}$ is provided the approach is clearly preferable.



(a) marginal posterior of μ_E (b) marginal posterior of σ_E

Figure 6: Sampled marginal posterior densities of μ_E and σ_E are shown, respectively. MCMC sampling was based on a blockwise random walk Metropolis algorithm. Here the true values $\mu_E = 15\,\mathrm{GPa}$ and $\sigma_E = 3\,\mathrm{GPa}$ have been recovered quite well.

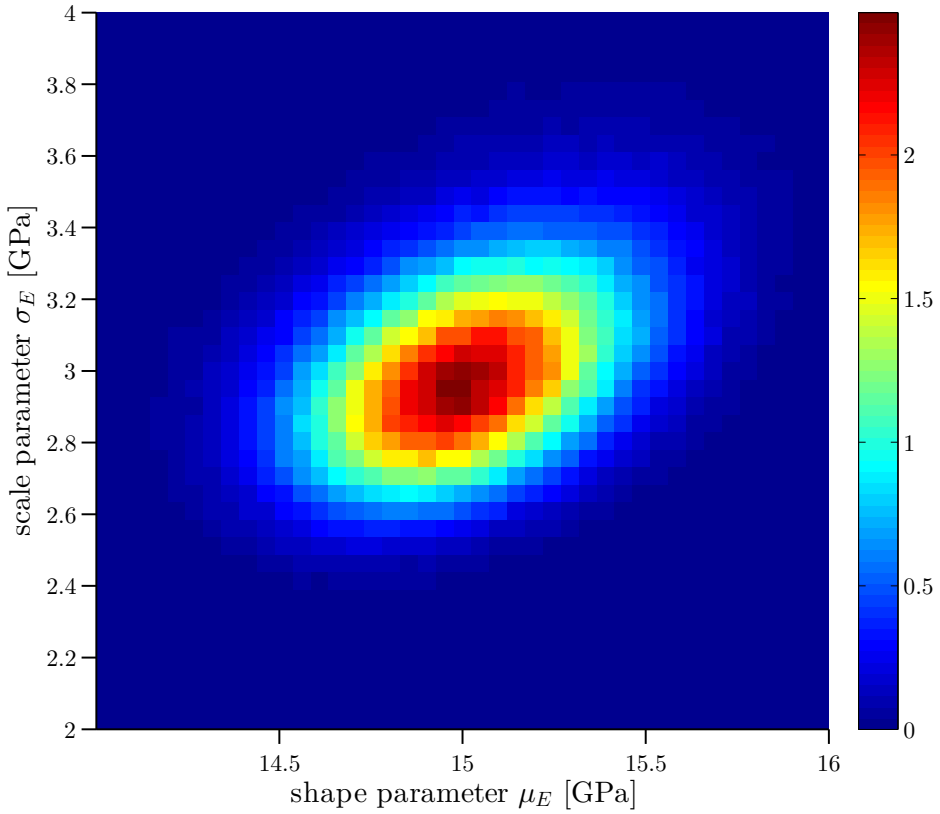


Figure 7: The sampled two-dimensional marginal posterior of the hyperparameter $\boldsymbol{\theta} = (\mu_E, \sigma_E)$ is shown. A blockwise random walk Metropolis MCMC algorithm was employed. Being priorly independent the components μ_E and σ_E are seen to be correlated a posteriori. The Pearson correlation coefficient amounts to $r_{\mu_E,\sigma_E} \approx 0.4$.

distribution. The influence of these insufficiently well-known inputs on the estimation of the quantities of interest and the computational expense this requires will be investigated in the future. The introduction of those inputs allows to better resemble realistic situations encountered in applied engineering practice.

Literature

[1] P. Barbillon, G. Celeux, A. Grimaud, Y. Lefebvre, and E. de Rocquigny. Nonlinear methods for inverse statistical problems. *Comput. Stat. Data Anal.*, 55(1):132 – 142, 2011.

[2] G. Celeux, A. Grimaud, Y. Lefèbvre, and E. de Rocquigny. Identifying intrinsic variability in multivariate systems through linearized inverse methods. *Inverse Prob. Sci. Engng.*, 18(3):401–415, 2010.

[3] E. de Rocquigny. *Modelling Under Risk and Uncertainty: An Introduction to Statistical, Phenomenological and Computational Methods*. Wiley, 1st edition, 2012.

[4] E. de Rocquigny and S. Cambier. Inverse probabilistic modelling of the sources of uncertainty: a non-parametric simulated-likelihood method with application to an industrial turbine vibration assessment. *Inverse Prob. Sci. and Eng.*, 17(7):937–959, 2009.

[5] C. Desceliers, R. Ghanem, and C. Soize. Maximum likelihood estimation of stochastic chaos representations from experimental data. *Int. J. Numer. Meth. Engng.*, 66:978–1001, 2006.

[6] S. Fu. *Inversion probabiliste bayésienne en analyse d'incertitude*. PhD thesis, Université Paris-Sud, Faculté des Sciences d'Orsay, 2012. in English.

[7] C.P. Robert and G. Casella. *Monte Carlo statistical methods (2^{nd} Ed.)*. Springer Series in Statistics. Springer Verlag, 2004.

[8] B. Sudret. Meta-models for structural reliability and uncertainty quantification. In K.K. Phoon, M. Beer, S.T. Quek, and S.D. Pang, editors, *Proc. 5th Asian-Pacific Symp. Struct. Reliab. (APSSRA'2012), Singapore*, pages 53–76, 2012. Keynote lecture.