

SOLUTION OF INVERSE PROBLEMS USING BAYESIAN APPROACH WITH APPLICATION TO ESTIMATION OF MATERIAL PARAMETERS IN DARCY FLOW

Simona DOMESOVA^{1,2,3}, Michal BERES^{1,2,3}

¹Department of Applied Mathematics, Faculty of Electrical Engineering and Computer Science, VSB–Technical university of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava, Czech Republic

²IT4Innovations National Supercomputing Center, VSB–Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava, Czech Republic

³Institute of Geonics of the Czech Academy of Sciences, Studentska 1768, 708 00 Ostrava, Czech Republic

simona.domesova@vsb.cz, michal.beres@vsb.cz

DOI: 10.15598/aeec.v15i2.2236

Abstract. *Standard numerical methods for solving inverse problems in partial differential equations do not reflect a possible inaccuracy in observed data. However, in real engineering applications we cannot avoid uncertainties caused by measurement errors. In the Bayesian approach every unknown or inaccurate value is treated as a random variable. This paper presents an application of the Bayesian inverse approach to the reconstruction of a porosity field as a parameter of the Darcy flow problem. However, this framework can be applied to a wide range of problems that involve some amount of uncertainty. Here the material field is modeled as a Gaussian random field, which is expressed as a function of several random variables. The information about these random variables is given by the resulting posterior distribution, which is then studied using the Cross-Entropy method and samples are generated using the Metropolis-Hastings algorithm.*

Keywords

Bayesian statistics, Cross-Entropy method, Darcy flow, Gaussian random field, inverse problem, Markov chain Monte Carlo methods, Metropolis-Hastings algorithm.

1. Introduction to the Bayesian Inversion

Consider a general inverse problem of finding a vector $u \in \mathbb{R}^n$ that satisfies an equation $y = G(u)$,

where $y \in \mathbb{R}^m$ is a known vector of observed data and $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a generally nonlinear function, usually called the observation operator. The inverse operator to G is unknown. In practical applications the observed data are corrupted by noise, therefore we consider a modified equation

$$y = G(u) + \eta, \tag{1}$$

where $\eta \in \mathbb{R}^m$ is the observational noise, i.e. we use a model with an additive noise. The resulting inverse problem of finding $u \in \mathbb{R}^n$ which satisfies Eq. (1) may have no exact solution or multiple solutions. However, the Bayesian approach naturally overcomes these difficulties.

In the Bayesian statistics the unknown vector u is understood as a random vector. Consequently, the objective is to find the joint probability density function (PDF) of this random vector, while y is given, i.e. to find the conditional probability density function $\pi(u|y)$, which specifies the so called posterior distribution. According to the Bayes theorem, the posterior distribution is defined by

$$\pi(u|y) = \frac{f(y|u) \pi_0(u)}{\int f(y|u) \pi_0(u) du}, \tag{2}$$

where $f(y|u)$ indicates the data likelihood, i.e. the PDF of y when u is given, and $\pi_0(u)$ indicates the prior distribution. In fact π_0 expresses our prior belief about the random vector u , which is then refined according to the observed data y . An example of the relation between these three probability density functions in one dimension is illustrated in Fig. 1.

At this point the Bayesian approach can be applied to the inverse problem given by Eq. (1). Let f_η de-

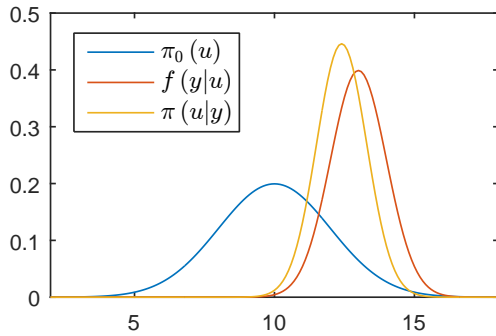


Fig. 1: Example of $\pi_0(u)$, $f(y|u)$ and $\pi(u|y)$ for $u \in \mathbb{R}$.

note the PDF of the observational noise η ; this random vector represents measurement errors, therefore we assume that its mean is a zero vector. Consequently, $f(y|u)$ equals the PDF of the noise η shifted by the vector $G(u)$, i.e. $f(y|u) = f_\eta(y - G(u))$ [2].

$$\pi(u|y) \propto f(y|u) \pi_0(u) = f_\eta(y - G(u)) \pi_0(u). \quad (3)$$

Further notice that the relation fully specifies the posterior distribution, therefore the aim of the following text is to find the posterior PDF only up to a normalizing constant.

2. The Gaussian Model

If the additive observational noise η represents a measurement error, it is natural to expect that it follows the normal distribution with zero mean, i.e. $\eta \sim \mathcal{N}(0, \Sigma)$, where $\Sigma \in \mathbb{R}^{m \times m}$ is the covariance matrix. For convenience we further assume that $\Sigma = \gamma^2 \mathbf{I}^{(m)}$, where $\mathbf{I}^{(m)} \in \mathbb{R}^{m \times m}$ is an identity matrix and $\gamma^2 \in \mathbb{R}_+$, i.e. the elements of the random vector η are independent. The data likelihood then satisfies:

$$f(y|u) = f_\eta(y - G(u)) = \frac{1}{\sqrt{(2\pi\gamma^2)^m}} \exp\left(-\frac{(y - G(u))^T (y - G(u))}{2\gamma^2}\right). \quad (4)$$

If the observational operator G is linear, this PDF is Gaussian, but in the general case it is not.

It is further assumed that the prior distribution of the random vector u is also Gaussian, for simplicity $u \sim \mathcal{N}(0, \sigma^2 \mathbf{I}^{(n)})$, where $\sigma^2 \in \mathbb{R}_+$. With this assumption:

$$\pi_0(u) = \frac{1}{\sqrt{(2\pi\sigma^2)^n}} \exp\left(-\frac{u^T u}{2\sigma^2}\right). \quad (5)$$

Consequently, the posterior PDF is given by:

$$\pi(u|y) \propto \exp\left(-\frac{(y - G(u))^T (y - G(u))}{2\gamma^2} - \frac{u^T u}{2\sigma^2}\right). \quad (6)$$

The use of the Gaussian prior information is well suited for porous media flow problems as the permeability is usually modeled using the lognormal distribution, which can be simply expressed as an exponential of the normal distribution. If we considered another prior distribution, the derivation of the posterior PDF would be similar.

If both data likelihood and prior information are Gaussian, the posterior distribution (as their product) is also Gaussian and it is fully specified by its mean and covariance operator. However, in this case a general nonlinear observation operator G is considered, therefore the posterior distribution is not Gaussian and its properties must be studied differently using appropriate techniques.

3. Properties of the Posterior Distribution

The next task is to gain as much information about the random parameters as possible from the resulting posterior PDF.

One way of studying the posterior distribution is through finding $\hat{u} \in \mathbb{R}^n$ that maximizes the posterior PDF. Denoting:

$$S(u) = \frac{(y - G(u))^T (y - G(u))}{2\gamma^2} + \frac{u^T u}{2\sigma^2}. \quad (7)$$

We can write:

$$\hat{u} = \arg \max_{u \in \mathbb{R}^n} \pi(u|y) = \arg \min_{u \in \mathbb{R}^n} S(u). \quad (8)$$

In the case of the Gaussian PDF, the value of \hat{u} equals to the mean. Generally \hat{u} is the mode, i.e. the “most probable” value of u , but it may not be uniquely determined.

By finding the approximation of the mode we only utilize a small portion of the information available about the unknown material parameters. In order to utilize all the information provided by the posterior distribution, we can use for example the Markov Chain Monte Carlo methods.

3.1. PDF Maximization Using the Cross-Entropy Method

The Cross-Entropy (CE) method is based on the Kullback-Leibler divergence minimization. Here the iterative CE algorithm, more specifically its variant for continuous optimization, is applied to the problem

$$\min_{u \in \mathbb{R}^n} S(u). \quad (9)$$

In this case, we are not interested in the minimal value but in the argument of the minimum. The Gaussian distribution $\mathcal{N}(\mu_0, \sigma_0^2 \mathbf{I}^{(n)})$ is chosen as an arbitrary distribution. The distribution parameters μ_0 and σ_0^2 are updated smoothly according to selected elite samples; it converges to a degenerated distribution with mean equal to \hat{u}_{CE} and zero variance. The value of \hat{u}_{CE} is the CE approximation of the mode (or one of the modes) \hat{u} . The derivation of the algorithm is thoroughly explained in [3].

The iterative algorithm of the CE method works with input parameters:

- $M \in \mathbb{N}$ (number of samples generated in each iteration),
- $\nu \in (0, 1)$ (parameter specifying the number of elite samples, typically between 0.8 and 0.99),
- $\alpha \in (0, 1)$ (smoothing parameter of the mean, typically between 0.7 and 1.0),
- $\beta \in (0, 1)$ (parameter of variance smoothing, typically between 0.8 and 0.99),
- $q \in \mathbb{N}$ (parameter of variance smoothing, typically between 5 and 10),
- $\mu_0 \in \mathbb{R}^n$ (initial mean),
- $\sigma_0^2 \in \mathbb{R}_+^n$ (initial variance).

And with the function $S : \mathbb{R}^n \rightarrow \mathbb{R}$ specified above. The steps of the algorithm are as follows:

- Set $t = 0$.
- Generate M independent samples $\{\mathbf{X}_1, \dots, \mathbf{X}_M\}$ from $\mathcal{N}(\mu_t, \sigma_t^2)$.
- Calculate $\mathbf{Y} = \{Y_1, \dots, Y_M\} = \{S(\mathbf{X}_1), \dots, S(\mathbf{X}_M)\}$.
- Calculate γ as a $(1 - \nu)$ -quantile of the set \mathbf{Y} .
- For $j \in \{1, \dots, n\}$ calculate the elements of the updated parameters μ_t and σ_t^2 using the formulas

$$(\mu_t)_j = \frac{\sum_{i=1}^M I_{\{S(\mathbf{x}_i) \geq \gamma\}} X_{ij}}{\sum_{i=1}^M I_{\{S(\mathbf{x}_i) \geq \gamma\}}}, \quad (11)$$

and

$$(\sigma_t^2)_j = \frac{\sum_{i=1}^M I_{\{S(\mathbf{x}_i) \geq \gamma\}} (X_{ij} - (\mu_t)_j)^2}{\sum_{i=1}^M I_{\{S(\mathbf{x}_i) \geq \gamma\}}}, \quad (12)$$

where X_{ij} denotes the j -th element of the vector \mathbf{X}_i . The operator I returns 1, when the statements in the brackets holds; otherwise, it returns 0.

- Set $\beta_t = \beta - \beta \cdot (1 - \frac{1}{t})^q$ and smooth the parameters using formulas

$$\mu_t = \alpha \cdot \mu_t + (1 - \alpha) \cdot \mu_{t-1}, \quad (13)$$

and

$$\sigma_t^2 = \beta_t \cdot \sigma_t^2 + (1 - \beta_t) \cdot \sigma_{t-1}^2. \quad (14)$$

- If the stopping condition is not fulfilled, increment the value of t by one and continue with the step 2.
- Return $\hat{u}_{CE} = \sigma_t$.

3.2. Sampling from $\pi(u|y)$ Using the Metropolis-Hastings Algorithm

The Metropolis-Hastings (MH) algorithm belongs to the Markov chain Monte Carlo methods, the principle of these methods is explained e.g. in [4]. The MH algorithm provides approximate samples from a given PDF known up to a normalizing constant, which is exactly this case. The method is particularly useful in the case of high-dimensional random vectors u , see [1]. We take:

$$q(u, x) = \frac{1}{\sqrt{(2\pi\sigma_{MH}^2)^n}} \exp\left(-\frac{(u-x)^T(u-x)}{2\sigma_{MH}^2}\right). \quad (15)$$

As the instrumental PDF and we can choose e.g. $u_0 = \hat{u}_{CE}$ or $u_0 = \mathbf{0}$ as the initial state, depending on the purpose of the simulation. The MH algorithm for sampling from the posterior PDF $\pi(u|y)$ reads as follows:

- for $t = 0, 1, \dots, T$
 - generate x_t from $q(u_t, x_t)$ and compute the acceptance probability:
 - $\alpha(u_t, x_t) = \min\left\{\frac{\pi(x_t|y)q(x_t, u_t)}{\pi(u_t|y)q(u_t, x_t)}, 1\right\}$,
 - set $u_{t+1} = x_t$ with probability $\alpha(u_t, x_t)$, otherwise set $u_{t+1} = u_t$.

The samples u_1, \dots, u_T are not independent, but u_T is approximately distributed according to $\pi(u|y)$ for high values of T .

We can further substitute Eq. (6) for $\pi(\cdot|\cdot)$ in Eq. (16) and notice that $q(u, x)$ is symmetric, we thus

$$\alpha(u_t, x_t) = \min \left\{ \exp \left(\frac{\|y - G(u_t)\|^2 - \|y - G(x_t)\|^2}{2\gamma^2} + \frac{\|u_t\|^2 - \|x_t\|^2}{2\sigma^2} \right), 1 \right\}. \tag{17}$$

obtain a simplified formula for the acceptance probability calculation Eq. (17).

The values $\|y - G(u_t)\|^2$ and $\|u_t\|^2$ are not calculated again, they are known from the previous iteration.

The acceptance rate $\bar{\alpha}$ is defined as the average of the acceptance probability $\alpha(u_t, x_t)$ over the iterations, i.e. see [8].

$$\bar{\alpha} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \alpha(u_t, x_t). \tag{18}$$

A high value of $\bar{\alpha}$ indicates a high correlation between consecutive samples and on the other hand a low value of $\bar{\alpha}$ means that a high number of proposed samples x_t is rejected. The value of the acceptance rate can be adjusted by the choice of the vector σ_{MH} .

4. Application to a Gaussian Random Field Reconstruction

In this paper, the Bayesian inverse approach is used for a reconstruction of a parameterized material field. This means that the structure of the material field is described as a function of certain number of random variables, grouped into a random vector. The Bayesian inversion is then used to determine the probability distribution of this random vector, according to observed data based on a specific sample of the material field. The principles are demonstrated on the Darcy flow problem, but they can be used generally for the reconstruction of parameters of different partial differential equations.

Consider a two field form of the groundwater flow boundary value problem on a unit square with a prescribed pressure on the left and the right side and no flow on the upper and the lower side:

$$\left. \begin{aligned} v + k\nabla p &= \mathbf{0} \\ \text{div}(v) &= 0 \end{aligned} \right\} \quad \text{in } D = \langle 0, 1 \rangle \times \langle 0, 1 \rangle \tag{19}$$

$$\begin{aligned} p &= 1 && \text{on } 0 \times \langle 0, 1 \rangle \\ p &= 0 && \text{on } 1 \times \langle 0, 1 \rangle \\ v \cdot n_0 &= 0 && \text{on } \langle 0, 1 \rangle \times \{0, 1\}, \end{aligned}$$

where $p : D \rightarrow \mathbb{R}$ is the pore pressure, $v : D \rightarrow \mathbb{R}^2$ is the Darcy's velocity, n_0 is the unit outer normal to the boundary, and k is the permeability of the material. We consider a non-homogeneous isotropic material and we use a model recommended in [7]. The permeability

is modeled as $k = \exp(\phi)$, where ϕ is the porosity, which is modeled as a Gaussian random field specified by zero mean and the autocovariance function Eq. (20) for details see [5].

$$c(x, y) = \rho^2 \exp \left(-\frac{\|x - y\|}{\lambda} \right). \tag{20}$$

4.1. The Direct Problem

The model problem Eq. (19) is solved by the Mixed Finite Element Method (MFEM) using the lowest order Raviart-Thomas discretization, as described in [6]. In the continuous case, the material was represented by a Gaussian random field. Here the input of the direct problem is a sample of the random vector $\tilde{\phi}$ (defined as a discretization of the Gaussian random field ϕ on a regular $N \times N$ grid). Outputs are the discretization \tilde{p} of the pore pressure field and the discretization $\tilde{v} = (\tilde{v}_1, \tilde{v}_2)$ of the Darcy's velocity vector field.

The elements of the covariance matrix $\mathbf{C} \in \mathbb{R}^{N^2 \times N^2}$ of the random vector $\tilde{\phi}$ are given by $\mathbf{C}_{ij} = c(g_i, g_j)$, where g_i and g_j are grid points. In [5] we use the Cholesky decomposition to generate the field, but for this purpose it is preferable to use the eigenvalue decomposition of \mathbf{C} :

$$\mathbf{C} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}, \tag{21}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with eigenvalues on the diagonal, sorted in descending order, and \mathbf{Q} is a matrix with eigenvectors in columns. If we assume $u \sim \mathcal{N}(0, \mathbf{I}_{N^2})$, then the realization of the random vector $\tilde{\phi}$ correlated according to \mathbf{C} is obtained as:

$$\tilde{\phi} = \mathbf{Q}\mathbf{\Lambda}^{\frac{1}{2}}u, \tag{22}$$

where $\tilde{\phi}$ is a sample of the discretized Gaussian random field. The random vector u can be understood as the input of the direct problem instead of $\tilde{\phi}$, because the relation between these vectors is deterministic. At this point we can define the Bayesian inverse problem.

4.2. The Inverse Problem

We assume that \tilde{p} , \tilde{v}_1 and \tilde{v}_2 are known, presumably only in several grid points. Let P , V_1 and V_2 denote the sets of points, where the pressure and the velocity (in certain directions), respectively, is known. Therefore the observation operator G is given by:

$$G(u) = (\tilde{p}|_P, \tilde{v}_1|_{V_1}, \tilde{v}_2|_{V_2}). \tag{23}$$

The application of this operator consists of transforming the vector u into a sample $\tilde{\phi}$ of the discretized random field according to Eq. (22), finding the approximate solution of the partial differential equation using the MFEM, and extracting the restricted vectors $\tilde{p}|_P$, $\tilde{v}_1|_{V_1}$ and $\tilde{v}_2|_{V_2}$. Again the observed data are corrupted by noise η . With this notation the Bayesian inverse problem is defined as finding the posterior distribution of the random vector $u \in \mathbb{R}^n$ (for now $n = N^2$) that satisfies the equation:

$$y = G(u) + \eta. \tag{24}$$

The posterior distribution is defined by Eq. (6), where $\sigma = 1$. The length of the random vector is N^2 (number of all grid points), therefore sampling from the posterior PDF by the MH algorithm or maximizing using the CE method is computationally intensive.

4.3. Dimension Reduction

The dimension of this inverse problem (number of unknown parameters) can be decreased by neglecting small eigenvalues and corresponding eigenvectors of the correlation matrix \mathbf{C} . This matrix is positive definite, its spectrum for $N = 50$, $\rho = 1$, $\lambda = 0.4$ is shown in Fig. 2. Therefore \mathbf{Q} is an orthogonal matrix and Eq. (22) can be rewritten as:

$$\tilde{\phi} = \sum_{i=1}^{N^2} \sqrt{\lambda_i} u_i q_i, \tag{25}$$

where λ_i is an eigenvalue and q_i is the corresponding eigenvector; $\lambda_i \geq \lambda_j$ for $i < j$. Figure 3 shows examples of eigenvectors q_{14} , q_{15} and q_{18} reshaped to the square domain D . By truncating this sum to $n < N^2$ summands, we obtain an approximation of $\tilde{\phi}$.

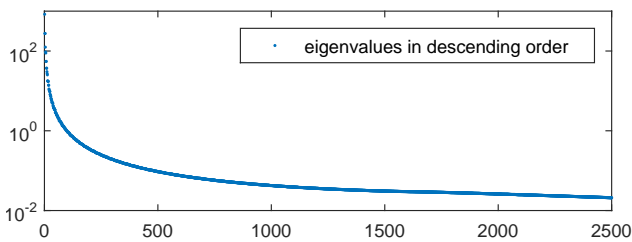


Fig. 2: Spectrum of \mathbf{C} for $N = 50$, $\rho = 1$, $\lambda = 0.4$.

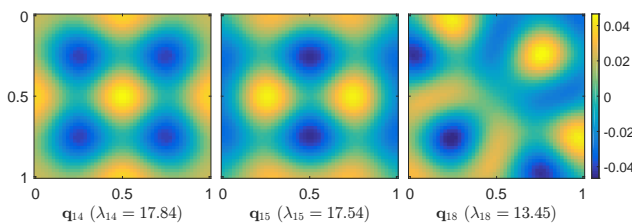


Fig. 3: Examples of eigenfunctions and corresp. eigenvalues.

Let $\mathbf{\Lambda}^{(n)} \in \mathbb{R}^{n \times n}$ and $\mathbf{Q}^{(n)} \in \mathbb{R}^{N^2 \times n}$ denote a diagonal matrix containing n largest eigenvalues of \mathbf{C} and a rectangular matrix containing the corresponding eigenvectors, respectively. The approximation $\tilde{\phi}^{(n)}$ of the Gaussian random field using n eigenvectors is now given by:

$$\tilde{\phi}^{(n)} = \sum_{i=1}^n \sqrt{\lambda_i} u_i q_i = \mathbf{Q}^{(n)} \left(\mathbf{\Lambda}^{(n)} \right)^{\frac{1}{2}} u, \tag{26}$$

where $u \sim \mathcal{N}(0, \mathbf{I}^{(n)})$. The dimension of the inverse problem is thus decreased to n . The value of $n \in \{1, \dots, N^2\}$ can be understood as a parameter of the observation operator G that also determines the length of the studied random vector, i.e. the number of random variables that are used to describe the randomness. Therefore the notation $G^{(n)}$ will be used for the observation operator in the rest of the text.

5. Numerical Experiments

For the numerical experiments the following settings were chosen: $N = 50$, $\rho = 1$, $\lambda = 0.4$, P is a set of 100 grid points at the Neumann boundary of the domain, V_1 is a set of 100 points regularly distributed over the domain, and V_2 is empty. The sets P and V_1 are illustrated by the red dots in Fig. 4. The observation operator $G^{(n)}$ is fully specified by these settings. At this moment it is necessary to obtain a vector of observed data y ; it will be simulated and it will serve as the vector of measurements (one of the inputs of the Bayesian inverse problem).

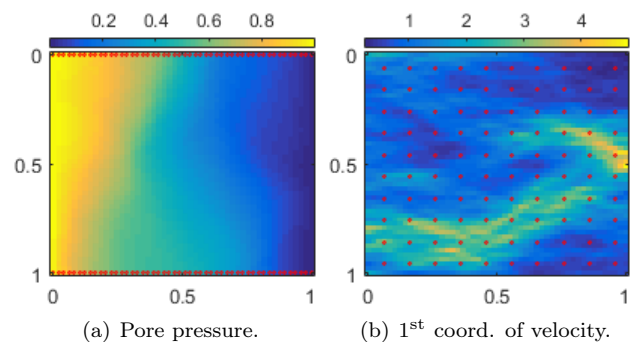


Fig. 4: The illustration of the measurements $\tilde{v}_1|_{V_1}$ and $\tilde{p}|_P$ before the application of the noise.

5.1. Simulated Observed Data

The covariance matrix was constructed as described in Subsec. 4.1. A sample of the discretized Gaussian random field was then generated according to Eq. (22). Figure 5 shows a realization of this random vector reshaped to the square domain D .

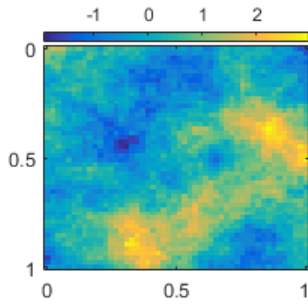


Fig. 5: Original porosity field $\mathbf{Q}\mathbf{\Lambda}^{\frac{1}{2}}u$ used to simulate a vector of observed data.

For this material field, the resulting vectors $\tilde{v}|_V$ and $\tilde{p}|_P$ were calculated using MFEM, see Fig. 4 (the pore pressure and the first coordinate of the velocity). To obtain the simulated observed data y , the output $(\tilde{v}|_V, \tilde{p}|_P)$ was corrupted by the additive Gaussian noise with parameter $\gamma = 0.05$. The values of the vector y are understood as noisy measurements and they are used for the calculation of the material parameters probability distribution.

5.2. Bayesian Inversion

By the substitution of the standard deviations γ and σ , the vector y , and the operator $G^{(n)}$ (for a given n) into Eq. (6), we obtain the posterior PDF of the random vector u . The dimension n of the problem can be chosen arbitrarily. For a comparison between the non-truncated and the truncated version of the problem, we choose values $N^2, 200$ and 20 , respectively.

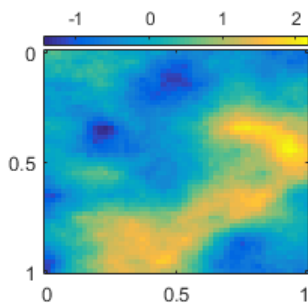


Fig. 6: Reconstructed material field calculated as $\mathbf{Q}\mathbf{\Lambda}^{\frac{1}{2}}\hat{u}_{CE}$.

First the material field was described using all 2500 eigenvectors of the matrix \mathbf{C} . The estimation $\hat{u}_{CE}^{(2500)}$ of the mode of the resulting posterior distribution was calculated by the CE method; 480 samples were generated in each of 10^3 iterations. The values of the remaining input parameters were set to $\nu = 0.975$ (this corresponds to 12 elite samples in each iteration), $\alpha = 0.99$, $\beta = 0.9$, $q = 8$, initial mean $\mu_0 = \mathbf{0}$, and initial variance $\sigma_0^2 = \mathbf{1}$. Figure 6 shows the reconstructed material field given by $\hat{u}_{CE}^{(2500)}$. Figure 7 shows the convergence of the

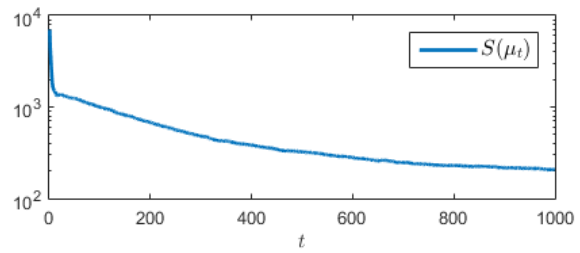


Fig. 7: Convergence of the CE method applied to $\hat{u}_{CE}^{(2500)}$ estimation.

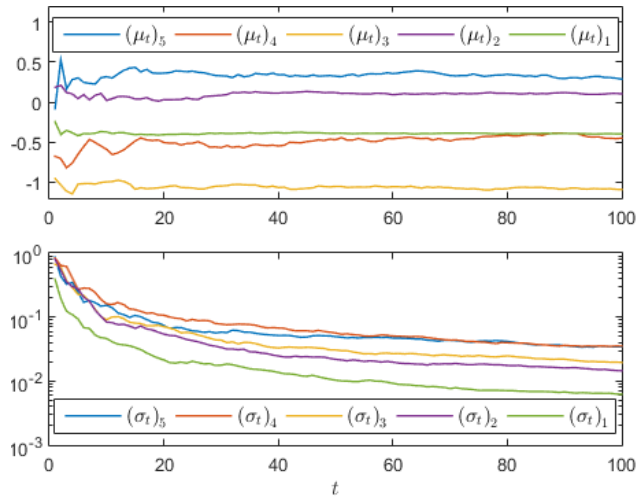


Fig. 8: Convergence of μ_t and σ_t elements corresponding to five highest eigenvalues.

CE method and the graphs in Fig. 8 show the convergence of the first 5 elements of μ_t and σ_t (corresponding to 5 highest eigenvalues of \mathbf{C}).

The reconstructed material fields in Fig. 9 use reduced number of eigenvectors for their description, specifically $n = 200$ and $n = 20$. The vectors $\hat{u}_{CE}^{(200)}$ and $\hat{u}_{CE}^{(20)}$ were calculated in 10^3 iterations of the CE method, 480 samples per iteration, the other input parameters of the CE method also remained unchanged.

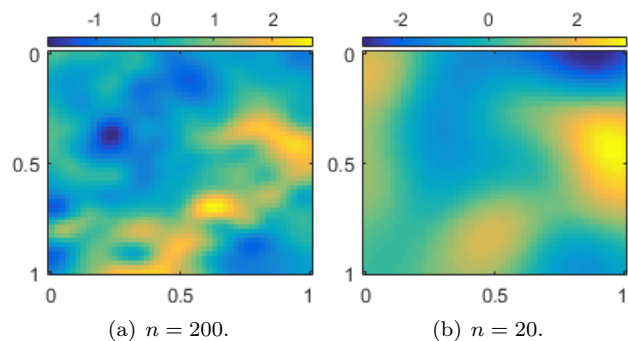


Fig. 9: Reconstructed material field $\mathbf{Q}(\mathbf{\Lambda}^{(n)})^{\frac{1}{2}}\hat{u}_{CE}^{(n)}$ with reduced number of random parameters.

The CE method was used only to estimate the mode of the posterior PDF, a more detailed description can be obtained by studying samples from the posterior distribution. For $n = 200$ and $n = 20$, approximate samples were obtained using the Metropolis-Hastings algorithm, see Subsec. 3.2. This algorithm was executed in three steps. In the initial run, 96 independent Markov chains of length $T = 1000$ were generated in parallel, each starting from $u_0 = \mathbf{0}$. The vectors u_T were used as initial states for the second run. During the second run of the MH algorithm, the parameter σ_{MH} of the instrumental density $q(u, x)$ was adjusted, the elements of σ_{MH} were updated adaptively to attain the acceptance rate close to $\bar{\alpha} = 0.25$ (respectively close to $\bar{\alpha} = 0.5$) for all elements of the vector u . The average of 96 vectors of σ_{MH} was used as the parameter of instrumental density for the final run. In the final run, the samples were again generated in parallel using 96 independent instances of the MH algorithm, each starting from a different initial state u_0 based on the preliminary runs.

In the case of $n = 20$, the adjustment of the elements of σ_{MH} (during the second run) was performed in batches of 100 consecutive states of the Markov chain. The acceptance rate for all of the 20 elements was estimated, and according to its values, the corresponding elements of the vector σ_{MH} were decreased or increased. The states of the Markov chain were generated element by element, therefore one “state” of the Markov chain was obtained in a loop over all the 20 elements. The graphs in Fig. 10 illustrate the changes of $(\sigma_{MH})_1$ and related changes of the acceptance rate of the first element of the random vector (corresponding to the high-

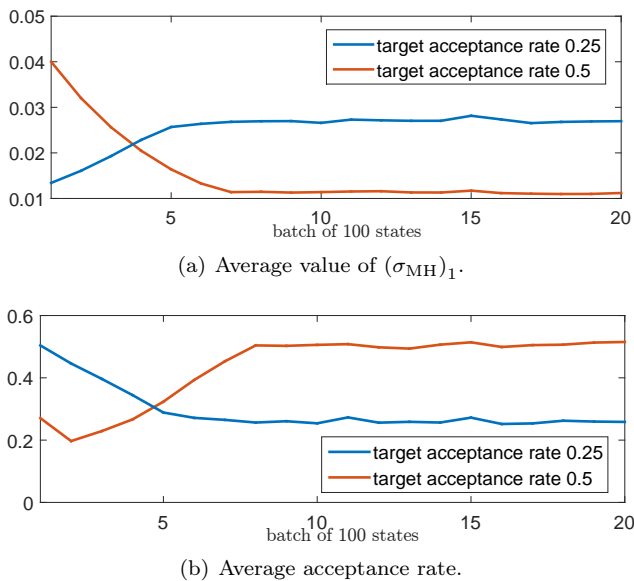


Fig. 10: Average values of $(\sigma_{MH})_1$ and corresponding acceptance rate during the preliminary run of the MH algorithm for $n = 20$.

est eigenvalue) during the second run. These graphs show the average over all of 96 instances for two target acceptance rates $\bar{\alpha} = 0.25$ and $\bar{\alpha} = 0.5$.

The final run was performed with σ_{MH} constant, 96 Markov chains of length 10000 were generated. The samples provided by the MH algorithm are correlated. To obtain nearly independent samples, the sample autocorrelation of the resulting Markov chains was analyzed. Based on the analysis, the variant “ $\bar{\alpha} = 0.5$ ” was chosen and every 1000th state was kept. This way the set $\mathbf{U}^{(20)} = \{u^{(20),1}, \dots, u^{(20),K}\}$ of $K = 960$ nearly uncorrelated samples from the posterior distribution was obtained. The states were transformed into the corresponding material fields using Eq. (26). Sample characteristics of this set $\mathbf{M}^{(20)} = \{m^{(20),1}, \dots, m^{(20),K}\} = \{\mathbf{Q}^{(20)} (\mathbf{\Lambda}^{(20)})^{\frac{1}{2}} u^{(20),1}, \dots, \mathbf{Q}^{(20)} (\mathbf{\Lambda}^{(20)})^{\frac{1}{2}} u^{(20),K}\}$ containing samples of the material field are shown in Fig. 11, the sample mean is a vector given by $\bar{m}^{(20)} = \frac{1}{K} \sum_{i=1}^K m^{(20),i}$ and the sample standard deviation is given by $\sqrt{\frac{1}{K-1} \sum_{i=1}^K (m^{(20),i} - \bar{m}^{(20)})^2}$.

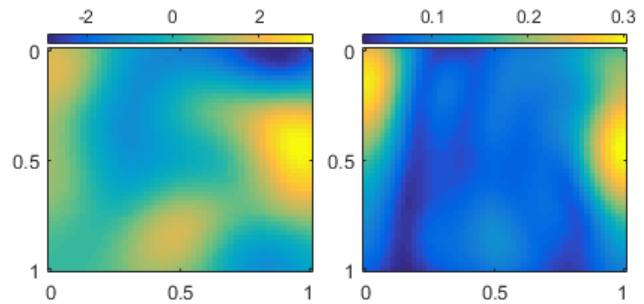


Fig. 11: Sample mean (left) and sample standard deviation (right) of the set $\mathbf{M}^{(20)}$.

In the case of $n = 200$, the procedure was the same except for a different batch size in the second run and a different length of Markov chains in the final run. The graphs in Fig. 12 show the average values of $(\sigma_{MH})_1$ and corresponding acceptance rates for consecutive batches of size 30. In the final run, 96 Markov chains of length 1200 were generated. Based on the analysis of the autocorrelation, the variant “ $\bar{\alpha} = 0.25$ ” was chosen and every 120th state was kept, i.e. the set $\mathbf{U}^{(200)}$ was obtained. The states were transformed into the corresponding material fields, the sample characteristics of the set $\mathbf{M}^{(200)}$ created this way are shown in Fig. 13. The sample standard deviation is lower around points, in which the velocity was measured. To illustrate this property, these points are drawn too.

Based on the samples $u^{(200),1}, \dots, u^{(200),K}$, the posterior distribution can be described in more detail. In Fig. 14 we can observe that the sample standard deviation of $(u)_i$ is close to 1 for higher values of i , i.e. close to the standard deviation of the prior PDF, while

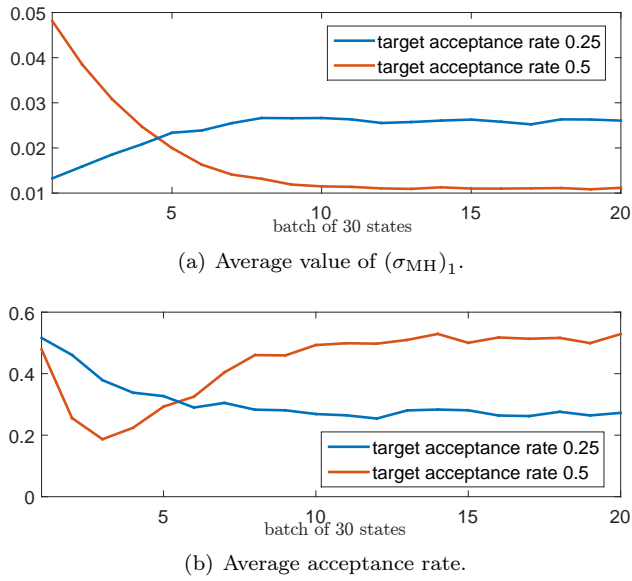


Fig. 12: Average values of $(\sigma_{MH})_1$ and corresponding acceptance rate during the preliminary run of the MH algorithm for $n = 200$.

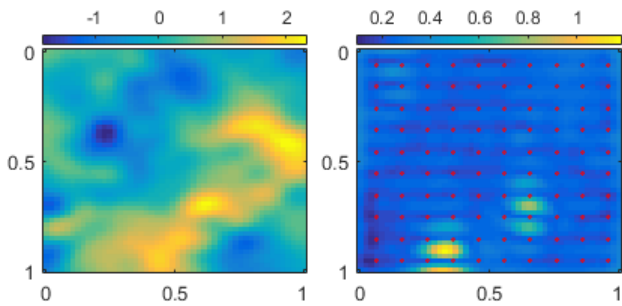


Fig. 13: Sample mean (left) and sample standard deviation (right) of the set $M^{(200)}$.

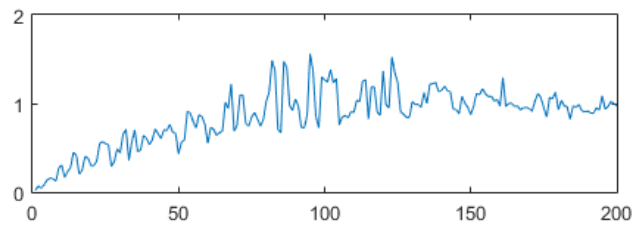


Fig. 14: Sample standard deviation of the set $U^{(200)}$.

the sample standard deviation of the components corresponding to high eigenvalues is lower.

The marginal distributions of the single components of the random vector u seem to be Gaussian, but the components are correlated. Figure 15 shows the absolute value of the sample correlation matrix. We can observe a strong correlation between some pairs of the random vector components. The pair of the components $(u)_{14}$ and $(u)_{15}$ and the pair $(u)_{18}$ and $(u)_{19}$ are the most correlated. This fact is illustrated in Fig. 16 and Fig. 17. The strong correlation in these pairs is

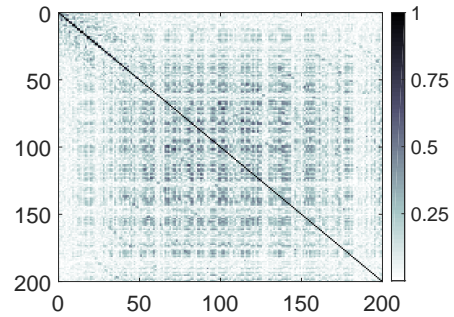


Fig. 15: Sample correlation matrix absolute value.

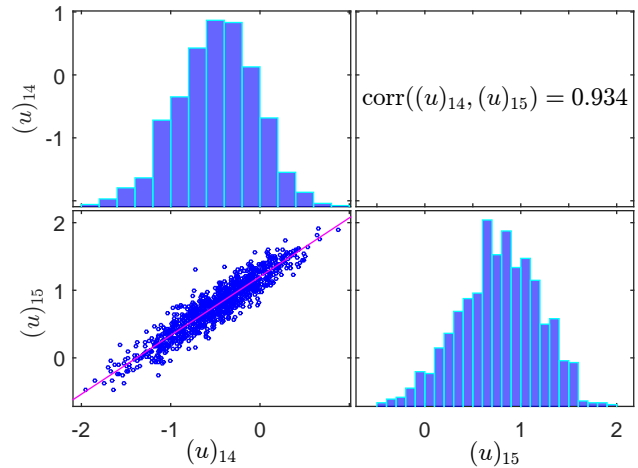


Fig. 16: Correlation between $(u)_{14}$ and $(u)_{15}$.

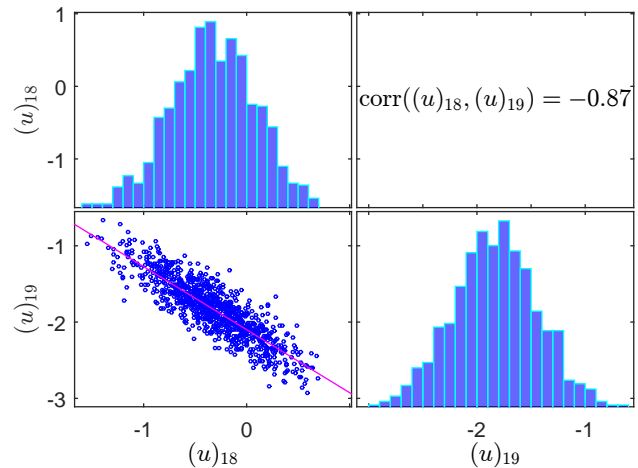


Fig. 17: Correlation between $(u)_{18}$ and $(u)_{19}$.

given by the specific character of the pairs of the corresponding eigenvectors. The eigenvectors q_{14} and q_{15} are illustrated in Fig. 3.

6. Conclusions

The paper presented the Bayesian inverse approach to material parameters estimation in the Darcy flow

model problem. The aim of this paper was to describe the posterior distribution of the material parameters. In the case of this model problem, the mode and the mean of the random material field matched and they corresponded to the original material field. Nearly uncorrelated samples from the posterior distribution were obtained using the MH algorithm. A lower efficiency of this way of sampling was given by the fact that the vast majority of the samples had to be neglected due to the strong autocorrelation. Using the samples, it was observed that the standard deviation of the random material field is influenced by the position of the domain points, where the measurements were taken.

The problem solved in this paper was specific due to the particular choice of the material in the form of the Gaussian random field. The framework can also be used for a description of different uncertain parameters, e.g. in the case of problems with given or parameterized material interfaces. The measurements and the amount of information hidden in them may also differ, mainly depending on the possibilities of a specific engineering application.

Each sample of the CE method and also each calculation of the acceptance probability during the MH algorithm involves a solution of the direct problem (i.e. an evaluation of the observation operator G). In this case it means a high number of the solutions of the Darcy flow problem using the MFEM [6]. Therefore the current work of the authors concentrates on alternative ways of solving the Darcy flow problem and other differential equation problems with random input data, e.g. using the stochastic collocation method and the stochastic Galerkin method. These methods require much lower computing time for the repeated solutions of the direct problem.

Acknowledgement

This work was supported by The Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPU II) project "IT4Innovations excellence in science - LQ1602". Work is partially supported by Grant of SGS No. SP2017/56, VSB-Technical University of Ostrava, Czech Republic.

References

- [1] STUART, A. M. Inverse problems: A Bayesian perspective. *Acta Numerica*. 2010, vol. 19, iss. 1, pp. 451–559. ISSN 0962-4929. DOI: 10.1017/S0962492910000061.
- [2] BUI-THANH, T., O. GHATTAS, J. MARTIN and G. STADLER. A computational framework for

infinite-dimensional bayesian inverse problems part I: The linearized case, with application to global seismic inversion. *SIAM Journal on Scientific Computing*. 2013, vol. 35, iss. 6, pp. 12494–12523. ISSN 1064-8275. DOI: 10.1137/12089586X.

- [3] RUBINSTEIN, R. Y. and D. P. KROESE. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. New York: Springer-Verlag, 2004. ISBN 978-1-4757-4321-0. DOI: 10.1007/978-1-4757-4321-0.
- [4] KROESE, D. P., T. TAIMRE and Z. I. BOTEV. *Handbook of Monte Carlo methods*. Hoboken: John Wiley & Sons, 2013. ISBN 978-0-470-17793-8. DOI: 10.1002/9781118014967.
- [5] BLAHETA, R., M. BERES and S. DOMESOVA. A study of stochastic FEM method for porous media flow problem. In: *1st International Conference on Applied Mathematics in Engineering and Reliability*. Boca Raton: CRC Press, 2016, pp. 281–289. ISBN 978-1-138-02928-6. DOI: 10.1201/b21348-47.
- [6] BLAHETA, R., M. HASAL, S. DOMESOVA and M. BERES. RT1-code: A mixed RT0-P0 Raviart-Thomas finite element implementation. *Institute of Geonics of the CAS* [online]. 2014. Available at: <http://www.ugn.cas.cz>.
- [7] HOEKSEMA, R. J. and P. K. KITANIDIS. Analysis of the Spatial Structure of Properties of Selected Aquifers. *Water resources research*. 1985, vol. 21, iss. 4, pp. 563–572. ISSN 0043-1397. DOI: 10.1029/WR021i004p00563.
- [8] ROBERT, C. and G. CASELLA. *Introducing Monte Carlo Methods with R*. 1st ed. New York: Springer-Verlag, 2010. ISBN 978-1-4419-1576-4. DOI: 10.1007/978-1-4419-1576-4.

About Authors

Michal BERES graduated from the Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava in 2015. He is currently a Ph.D. student at the same university. His research interests include uncertainty in the numerical models.

Simona DOMESOVA graduated from VSB-Technical University of Ostrava, where she studied Computational Mathematics, in 2015. Currently, she is a Ph.D. student at the same university. Her research interests include stochastic modeling.