L aboratório
N acional de
C omputação
C ientífica

COMOD
Coordenação de Modelagem Computacional

Relatório de Atividades

Servidor: Marcio Borges

Petrópolis-RJ
July 5, 2022

1 Method for generating permeability fields

Due to incomplete data about rock properties that show variability at multiple length scales, input parameters such as the permeability field, $\kappa(\mathbf{x}, \omega)$, are treated as random space functions with statistics inferred from geostatistical models (here $\mathbf{x} = (x_1, x_2, x_3)^{\mathsf{T}} \in \mathbb{R}^3$ and ω is a random element in the probability space). In line with Dagan (1989) and Gelhar (1993) the permeability field is modeled as a log-normally distributed random space function

$$\kappa(\mathbf{x}, \omega) = \exp\left[\mathsf{Y}(\mathbf{x}, \omega)\right], \tag{1}$$

where $\mathsf{Y}(\mathbf{x}, \omega) \sim \mathbb{N}(\mu_{\mathsf{Y}}, \mathcal{C}_{\mathsf{Y}})$ is a Gaussian random field characterized by its mean $\mu_{\mathsf{Y}} = \langle \mathsf{Y} \rangle$ and two-point covariance function

$$\mathcal{C}_{\mathsf{Y}}(\mathbf{x}, \mathbf{y}) = \mathsf{Cov}\left(\mathsf{Y}(\mathbf{x}), \mathsf{Y}(\mathbf{y})\right) = \mathsf{E}\Big[\left(\mathsf{Y}(\mathbf{x}) - \langle \mathsf{Y}(\mathbf{x}) \rangle\right)\left(\mathsf{Y}(\mathbf{y}) - \langle \mathsf{Y}(\mathbf{y}) \rangle\right)\Big]. \tag{2}$$

The Gaussian field Y can be represented as a series expansion involving a complete set of deterministic functions with correspondent random coefficients using the Karhunen-Loève (KL) expansion, proposed independently by Karhunen (1946) and Loève (1955). It is based on the eigen-decomposition of the covariance function. Depending on how fast the eigenvalues decay one may be able to retain only a small number of terms in a truncated expansion and, consequently, this procedure may reduce the search to a smaller parameter space. In uncertainty quantification methods for porous media flows, the KL expansion has been widely used to reduce the number of parameters used to represent the permeability field (Efendiev et al., 2005, 2006; Das et al., 2010; Mondal et al., 2010; Ginting et al., 2011, 2012). Another advantage of KL expansion lies on the fact that it provides orthogonal deterministic basis functions and uncorrelated random coefficients, allowing for the optimal encapsulation of the information contained in the random process into a set of discrete uncorrelated random variables (Ghanem and Spanos, 1991). This remarkable feature can be used to simplify the Metropolis-Hastings McMC Algorithm in the sense of the search may be performed in the space of discrete uncorrelated random variables ($\boldsymbol{\theta}$), no longer in the space of permeabilities which have a more complex statistical structure.

Here we recall the basic facts about the Karhunen-Loève expansion. Consider a random field $\mathsf{Y}(\mathbf{x}, \omega)$ defined on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ composed by the sample space, the ensemble of events and a probability measure, respectively, and indexed on a bounded domain $\mathcal{D} \in \mathbb{R}^3$. The process Y can be expressed as

$$\mathsf{Y}(\mathbf{x}, \omega) = \langle \mathsf{Y}(\mathbf{x}) \rangle + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\phi_i(\mathbf{x})\theta_i(\omega), \tag{3}$$

where λ_i and ϕ_i are the eigenvalues and eigenfunctions of the covariance function $\mathcal{C}_{\mathsf{Y}}(\mathbf{x}, \mathbf{y})$, respectively. By definition, $\mathcal{C}_{\mathsf{Y}}(\mathbf{x}, \mathbf{y})$ is bounded, symmetric and positive definite and has the following eigen-decomposition:

$$\mathcal{C}_{\mathsf{Y}}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{\infty} \lambda_i \phi_i(\mathbf{x})\phi_i(\mathbf{y}).$$

The eigenvalues and eigenfunctions of Eq. (3) are the solution of the homogeneous Fredholm integral equation of second kind given by

$$\int_{\mathcal{D}} \mathcal{C}_{\mathsf{Y}}(\mathbf{x},\mathbf{y})\phi(\mathbf{x})d\mathbf{x} = \lambda\phi(\mathbf{y}). \tag{4}$$

The solution of Eq. (4) forms a complete set of a square-integrable orthogonal eigenfunctions that satisfy the equation

$$\int_{\mathcal{D}} \phi_i(\mathbf{x})\phi_j(\mathbf{x}) = \delta_{ij},$$

in which δ_{ij} is the Kronecker-delta function. $\theta_i(\omega)$ is a set of independent random variables which can be expressed as

$$\theta_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{D}} \tilde{\mathsf{Y}}\phi_i(\mathbf{x})d\mathbf{x},$$

where $\tilde{\mathsf{Y}} = \mathsf{Y} - \langle \mathsf{Y} \rangle$ is the fluctuation. For practical implementations of the KL expansion the eigenvalues are arranged from the largest to smallest and the series is approximated by a finite number of terms, say the first m, giving

$$\mathsf{Y}(\mathbf{x},\omega) \approx \langle \mathsf{Y}(\mathbf{x}) \rangle + \sum_{i=1}^{\mathsf{m}} \sqrt{\lambda_i}\phi_i(\mathbf{x})\theta_i(\omega). \tag{5}$$

The corresponding covariance function is given by

$$\mathcal{C}_{\hat{Y}}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{\mathsf{m}} \lambda_i \phi_i(\mathbf{x})\phi_i(\mathbf{y}).$$

The factors affecting the convergence of the Karhunen-Loève series are the ratio of the length of the process over correlation parameter, the form of the covariance function, and the solution method for the eigensolutions of the covariance function (see Huang et al. (2001)). Next we discuss the field conditioning using the KL expansion.

1.1 Conditioning procedure

From measurements obtained in samples, the permeability field is known at sparse locations and this information can be incorporated in the *a priori* distribution.

Suppose that the Gaussian field Y (as well as the permeability field κ of Eq. (1)) is known at Nc positions $\{\mathring{\mathbf{x}}_i\}_{i=1}^{\mathsf{Nc}}$, that is:

$$\mathsf{Y}\left(\mathring{\mathbf{x}}_i\right) = \mathsf{Y}_i, \quad i = 1, \ldots, \mathsf{Nc}. \tag{6}$$

1.1 Conditioning procedure

To honor these measures in the field generation process with the KL expansion we consider the projection of random vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots]^{\mathsf{T}}$ (Eq. (3)) onto an appropriate subspace as proposed by Govinder et al. (2014). In their method the original $\boldsymbol{\theta}$ is replaced by its projection $\widetilde{\boldsymbol{\theta}}$ that has the same distribution as the original. Next, we briefly describe the method for the KL expansion truncated at m term.

Let $\mathbf{S} \in \mathbb{R}^{\mathsf{Nc}\times\mathsf{Nc}}$ be the covariance matrix of the observed locations

$$\mathbf{S}_{k,j} \equiv \mathcal{C}_{\mathsf{Y}}\left(\mathring{\mathbf{x}}_k, \mathring{\mathbf{x}}_j\right) = \sum_{i=1}^{\mathsf{m}} \lambda_i \phi_i(\mathring{\mathbf{x}}_k)\phi_i(\mathring{\mathbf{x}}_j), \quad k, j = 1, \ldots, \mathsf{Nc}. \tag{7}$$

Next, define $\mathbf{R}^{\mathsf{m}\times\mathsf{Nc}}$ to be the matrix with the columns given by the eigenfunctions at the observed locations $\{\mathring{\mathbf{x}}_i\}_{i=1}^{\mathsf{Nc}}$, then

$$\mathbf{R} \equiv \Big[\phi_i(\mathring{\mathbf{x}}_1), \phi_i(\mathring{\mathbf{x}}_2), \ldots, \phi_i(\mathring{\mathbf{x}}_{\mathsf{Nc}})\Big]. \tag{8}$$

By the last definition and making $\boldsymbol{\Lambda}$ the diagonal matrix of eigenvalues λ_i ($i = 1, \ldots, \mathsf{m}$), $\mathbf{S}$ can be rewritten as

$$\mathbf{S} = \mathbf{R}^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{R}.$$

To compute the mean $\widetilde{\mu_\theta}$ and covariance $\mathbf{M} = (m_{i,k})$ of the sequence of random variables $\{\theta_i\}_{i=1}^{\mathsf{m}}$ conditioned on $\mathring{\mathbf{X}} = \{\mathring{\mathbf{x}}_i\}_{i=1}^{\mathsf{Nc}}$ an usual formula for conditional means and covariance of Gaussian variables is used (Tong, 1990):

$$\widetilde{\mu_{\theta,i}} = \mathsf{E}\left[\theta_i | \mathring{\mathbf{X}}\right] = \sqrt{\lambda_i}\phi_i(\mathring{\mathbf{x}})\mathbf{S}^{-1}\left[\mathsf{Y}(\mathring{\mathbf{x}}) - \langle \mathsf{Y}(\mathring{\mathbf{x}})\rangle\right] \tag{9}$$

and

$$\begin{aligned} m_{i,j} &= \mathsf{Cov}\left(\theta_i, \theta_j | \mathring{\mathbf{X}}\right) \\ &= \delta_{i,j} - \sqrt{\lambda_j}\phi_j(\mathring{\mathbf{x}})\mathbf{S}^{-1}\phi_i(\mathring{\mathbf{x}})\sqrt{\lambda_i}. \end{aligned} \tag{10}$$

Thus, the projection matrix is given by

$$\mathbf{M} = \mathbf{I} - \boldsymbol{\Lambda}^{1/2}\, \mathbf{R}\, \mathbf{S}^{-1}\, \mathbf{R}^{\mathsf{T}}\, \boldsymbol{\Lambda}^{1/2}, \tag{11}$$

where I is the identity matrix. Finally the vector $\boldsymbol{\theta}$ conditioned on $\mathring{\mathbf{X}}$ is given by

$$\widetilde{\boldsymbol{\theta}} = \widetilde{\mu_{\boldsymbol{\theta}}} + \mathbf{M}\boldsymbol{\theta}. \tag{12}$$

Here $\mathbf{M}$ projects $\boldsymbol{\theta}$ onto the subspace that gives the $\widetilde{\mathsf{Y}}$ process conditional variance 0 at the locations $\{\mathring{\mathbf{x}}_i\}_{i=1}^{\mathsf{Nc}}$. The conditional representation $\widetilde{\mathsf{Y}}(\mathbf{x})$ of $\mathsf{Y}(\mathbf{x})$ reads as

$$\widetilde{\mathsf{Y}}(\mathbf{x},\omega) \approx \langle \mathsf{Y}(\mathbf{x})\rangle + \sum_{i=1}^{\mathsf{m}} \sqrt{\lambda_i}\phi_i(\mathbf{x})\widetilde{\theta}_i(\omega), \quad \mathbf{x} \in \mathcal{D}. \tag{13}$$

The matrix $\mathbf{M}$ projects the m dimensional vector $\boldsymbol{\theta}$ onto the $\mathsf{m} - \mathsf{Nc}$ dimensional subspace. Then, in order to have a sufficient number of degrees of freedom for the projection, we must have $\mathsf{m} > \mathsf{Nc}$

The number of terms used in the series can be chosen based on the energy represented by the sum of the eigenvalues. Then we define the relative energy for n terms as

$$\mathsf{ER_n} = \frac{\sum_{i=1}^{n} \lambda_i}{\sum_{j=1}^{m\to\infty} \lambda_j}. \tag{14}$$

2 NUMERICAL EXPERIMENTS

To test the implemented code, numerical experiments were performed, in which sets of 5000 fields (conditioned and unconditioned). Different covariance functions ($\mathcal{C}_{\mathsf{Y}}(\mathbf{x},\mathbf{y})$), correlation lengths ($\ell$) and number of expansion terms (m) were used in the tests. After obtaining the fields, the mean and covariance of the sets were computed and compared with the expected values.

In all cases, three-dimensional fields were generated with dimensions: $1.0 \times 0.6 \times 0.4$ and mesh $25 \times 15 \times 10$.

2.1 Squared Exponential Covariance

The two-point squared exponentially decaying covariance function is given by

$$\mathcal{C}_{\mathsf{Y}}(\mathbf{x},\mathbf{y}) = \sigma_{\mathsf{Y}}^2 \exp\left(-\frac{\|x_1-y_1\|^2}{\ell_1^2} - \frac{\|x_2-y_2\|^2}{\ell_2^2} - \frac{\|x_3-y_3\|^2}{\ell_3^2}\right), \tag{15}$$

with σ_{Y}^2 denoting the variance and $\ell_i > 0,\ i = 1,2,3$ the correlation length in each cartesian direction. More specifically, in this study, $\ell_1 = 0.25$, $\ell_2 = 0.20$ and $\ell_3 = 0.15$.

2.1 Squared Exponential Covariance

2.1.1 Unconditioned

Two realizations of the fields are displayed in Fig. 1.

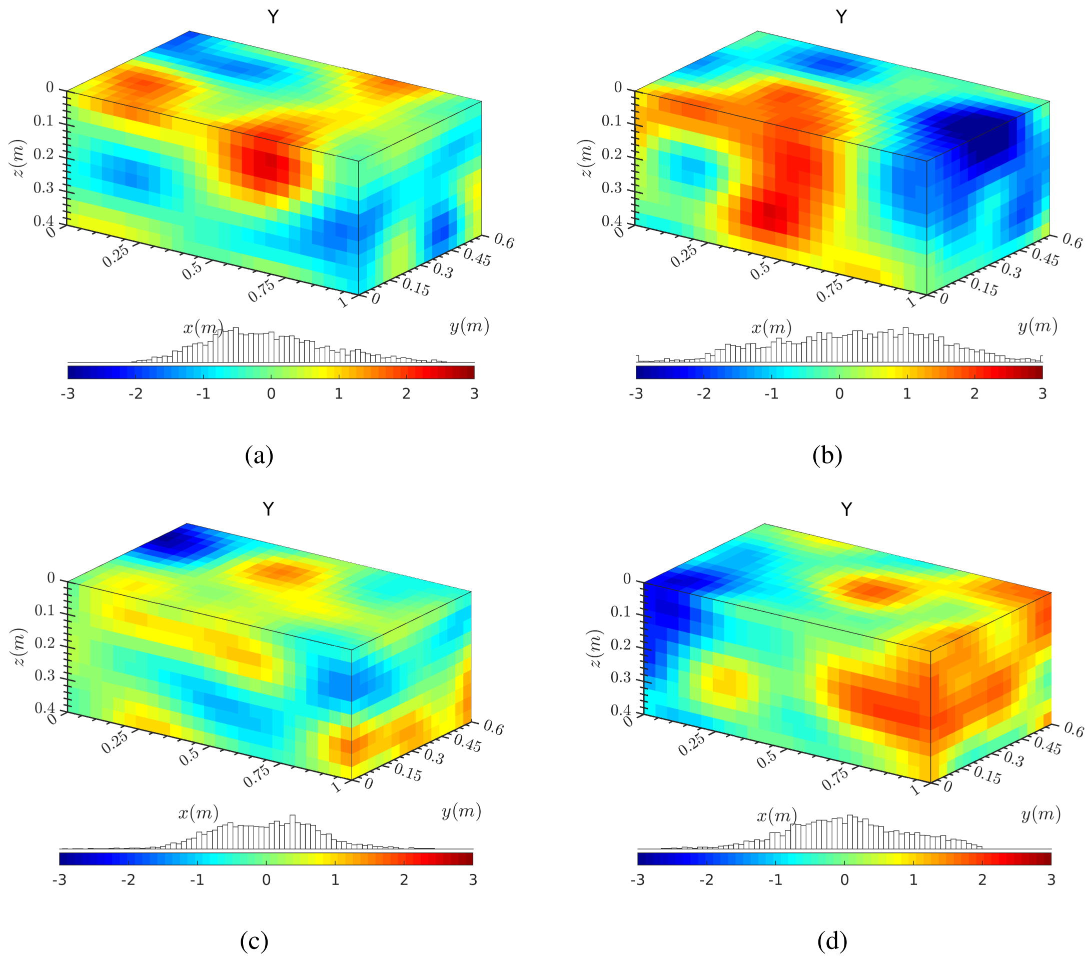
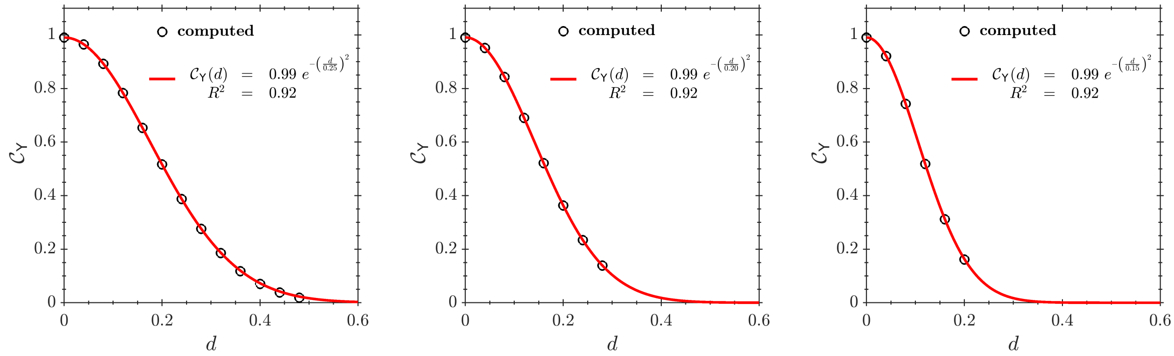


Figure 1: Realizations of random fields with squared exponential covariance (Eq. (17)).

Fig. 4 shows the computed covariance function obtained with 2000 realizations. The models obtained (red line) by fitting the computed points are in excellent agreement with the theoretical model (Eq. (17)).

2.1 Squared Exponential Covariance



(a) Direction x (b) Direction y (c) Direction z

Figure 2: Covariance as function of the lag distance $d = \|\mathbf{x} - \mathbf{y}\|$, computed with an ensamble of 2000 unconditioned fields.

2.1.2 Conditioned

In this section, an experiment similar to the previous one is conducted, however, the fields are conditioned at four points, as shown in the Tab. 1.

Table 1: Conditioned points

Point $\mathbf{x}$			
x	y	z	$Y(\mathbf{x})$
0.125	0.025	0.375	1.000
0.575	0.025	0.125	1.000
0.975	0.175	0.225	-1.000
0.975	0.525	0.175	-1.000

In Fig. 3 two realizations of the fields are presented and the cubes show the conditioned points.

2.1 Squared Exponential Covariance

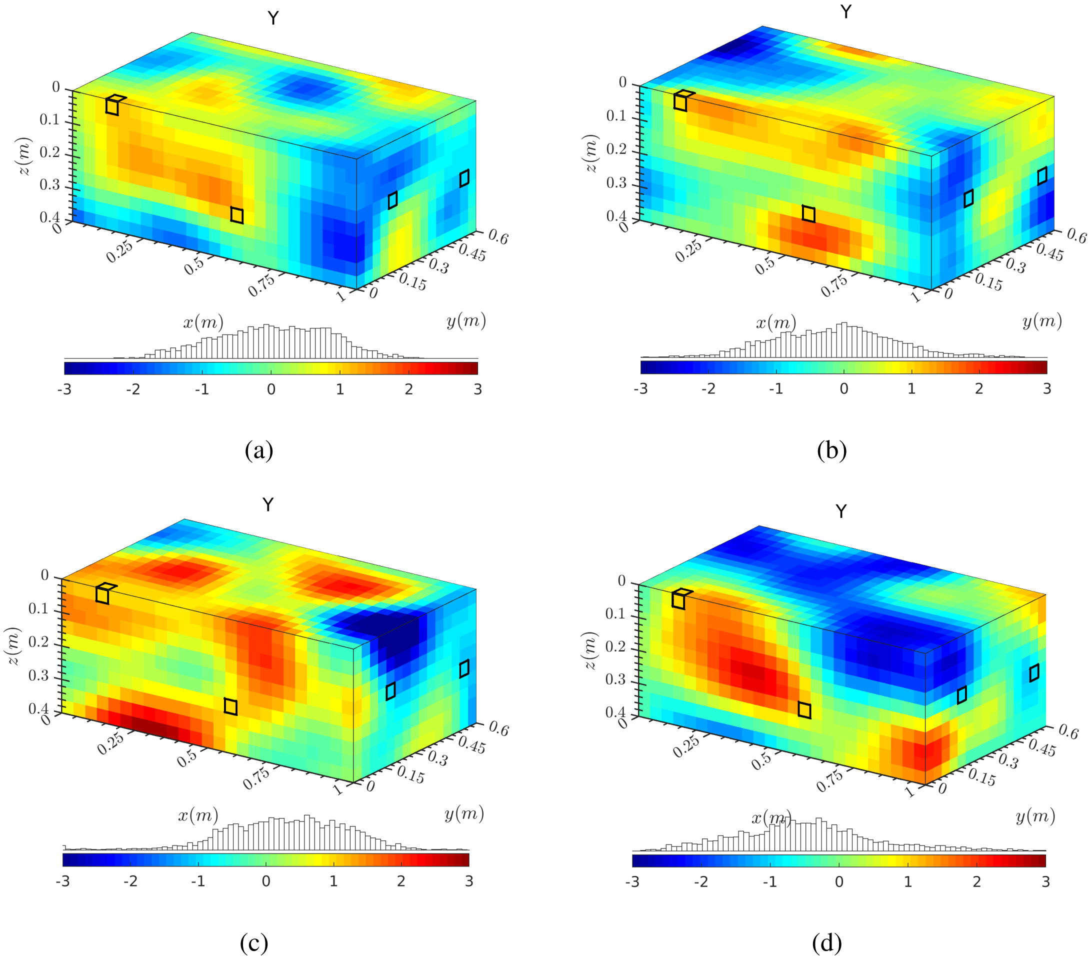
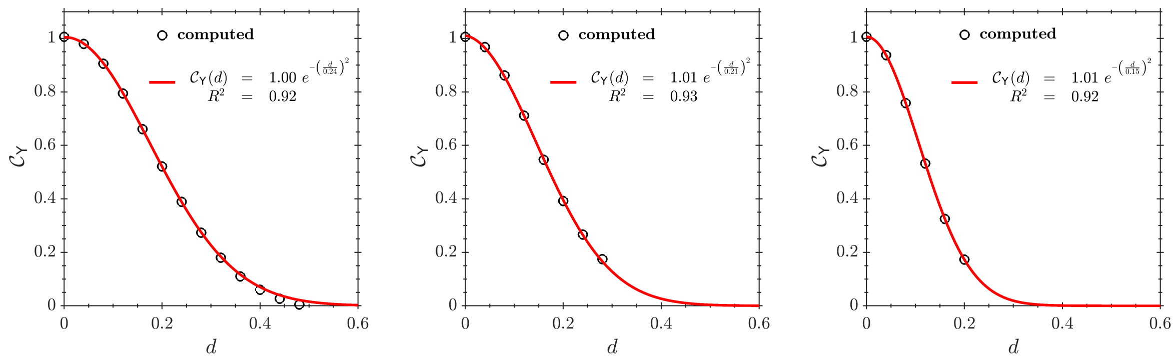


Figure 3: Realizations of random fields with squared exponential covariance (Eq. (17)).

Fig. 4 shows the computed covariance function obtained with 2000 realizations. The models obtained (red line) by fitting the computed points are in excellent agreement with the theoretical model (Eq. (17)).



(a) Direction x (b) Direction y (c) Direction z

Figure 4: Covariance as function of the lag distance $d = \|\mathbf{x} - \mathbf{y}\|$, computed with an ensamble of 2000 unconditioned fields.

3 Synthetic experiment

We consider three-dimensional fields with dimensions: $36.0 \times 36.0 \times 50.0$ $[mm \times mm \times mm]$ and mesh $36 \times 36 \times 50$. Two kind of two-point covariance functions are considered:

- Exponentially decaying covariance

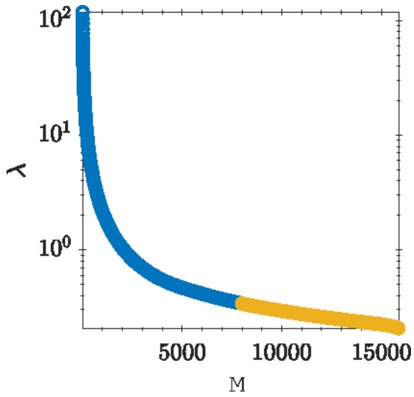
$$\mathcal{C}_Y(\mathbf{x}, \mathbf{y}) = \sigma_Y^2 \exp\left(-\frac{\|x_1 - y_1\|}{\ell_1} - \frac{\|x_2 - y_2\|}{\ell_2} - \frac{\|x_3 - y_3\|}{\ell_3} \right), \tag{16}$$

- Squared exponentially decaying covariance

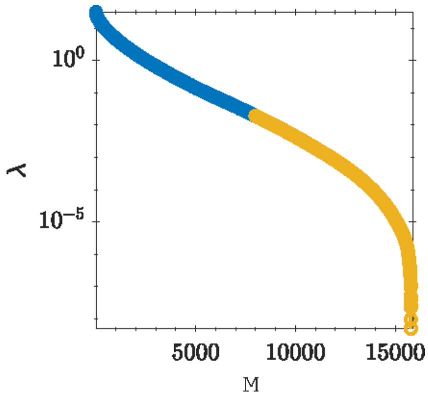
$$\mathcal{C}_Y(\mathbf{x}, \mathbf{y}) = \sigma_Y^2 \exp\left(-\frac{\|x_1 - y_1\|^2}{\ell_1^2} - \frac{\|x_2 - y_2\|^2}{\ell_2^2} - \frac{\|x_3 - y_3\|^2}{\ell_3^2} \right), \tag{17}$$

with σ_Y^2 denoting the variance and $\ell_i > 0,\ i = 1, 2, 3$ the correlation length in each cartesian direction. The correlation lengths used in this study are: $\ell_1 = \ell_2 = \ell_3 = 3.0mm$.

Samples of the fields generated with KL expansion are displayed in Figs. 6 - 7



(a) Exponential



(b) Square exponential

Figure 5: Eigenvalues of the covariance (Eq. (3)) as function of the number of terms in descending order.

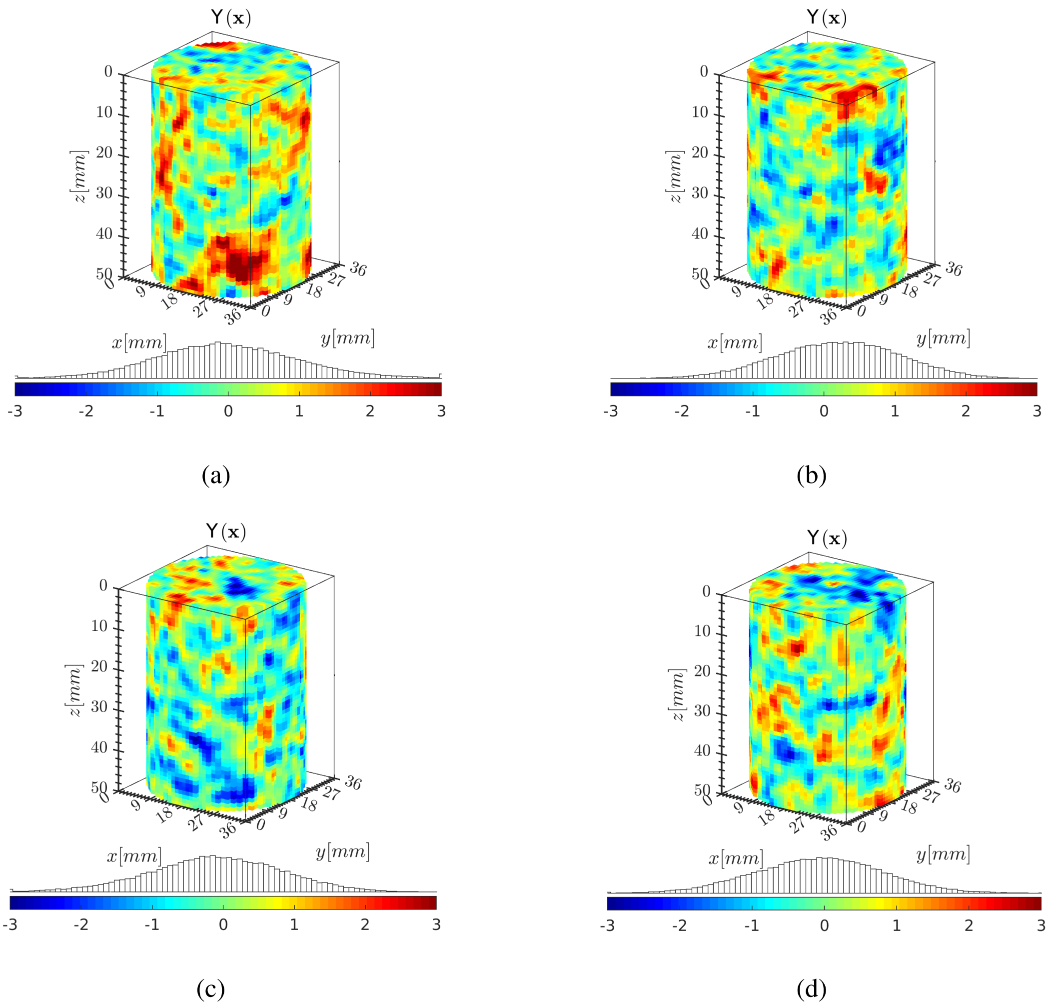
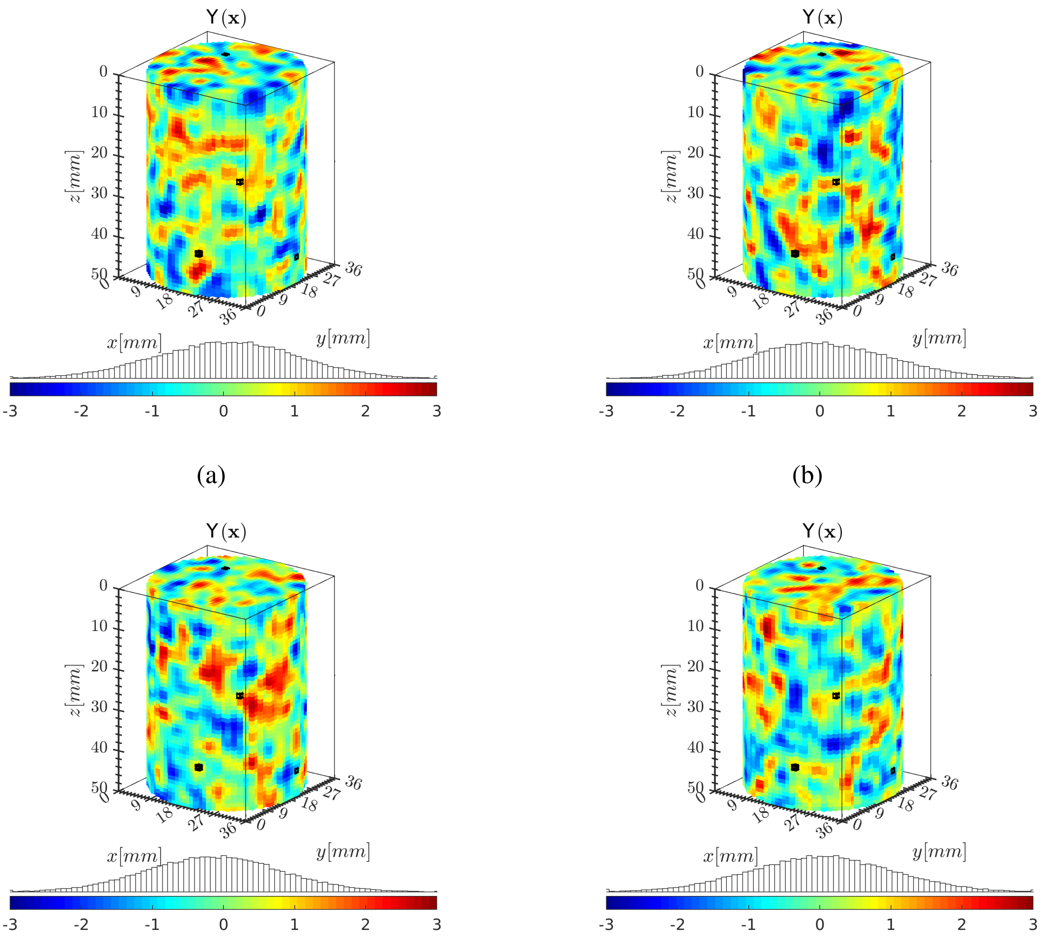


Figure 6: Samples of random fields with exponentially decay covariance.



(a) (b)

(c) (d)

Figure 7: Samples of random fields with exponentially decay covariance.

REFERENCES

G. Dagan. *Flow and transport in porous formations*. Springer-Verlag, 1989.

N. N. Das, B. P. Mohanty, and Y. Efendiev. Characterization of effective saturated hydraulic conductivity in an agricultural field using Karhunen-Loève expansion with the Markov chain Monte Carlo technique. *Water Resources Research*, 46, W06521 2010. doi: 10.1029/2008WR007100.

Y. Efendiev, A. Datta-gupta, V. Ginting, X. Ma, and B. Mallick. An efficient two-stage Markov chain Monte Carlo method for dynamic data integration. *Water Resources Research*, 41: 12423, 2005.

Y. Efendiev, T. Hou, and W. Luo. Preconditioning Markov Chain Monte Carlo Simulations Using Coarse-Scale Models. *SIAM J. Sci. Comput.*, 28:776–803, 2006.

L. W. Gelhar. *Stochastic subsurface hydrology*. Englewood Cliffs. Prentice-Hall, 1993.

R. Ghanem and P.D. Spanos. *Stochastic Finite Element: A Spectral Approach*. Springer, New York, 1991.

V. Ginting, F. Pereira, M. Presho, and S. Wo. Application of the two-stage Markov chain Monte Carlo method for characterization of fractured reservoirs using a surrogate flow model. *Computational Geosciences*, 15:691–707, 2011.

V. Ginting, F. Pereira, and A. Rahunanthan. Multiple Markov chains Monte Carlo approach for flow forecasting in porous media. *Procedia Computer Science*, 5:707–716, 2012.

K. S. Govinder, M. E. Ossiander, M. Peszynska, and V. S. Vasylkivska. Conditional stochastic simulations of flow and transport with karhunen-loève expansions, stochastic collocation, and sequential gaussian simulation. *Journal of Applied Mathematics*, 2014.

S. P. Huang, S. T. Quek, and K. K. Phoon. Convergence study of the truncated Karhunen-Loève expansion for simulation of stochastic processes. *International Journal for Numerical Methods in Engineering*, 52(9):1029–1043, 2001.

K. Karhunen. Zur spektraltheorie stochastischer prozesse. *Ann. Acad. Sci. Fennicae*, 1946.

M. M. Loève. *Probability Theory*. Princeton, N.J., 1955.

A. Mondal, Y. Efendiev, B. Mallick, and A. Datta-Gupta. Bayesian uncertainty quantification for flows in heterogeneous porous media using reversible jump Markov chain Monte Carlo methods. *Advances in Water Resources*, 33(3):241 – 256, 2010.

Y.L. Tong. *The Multivariate Normal Distribution*. Applications of Mathematics. Springer-Verlag, 1990. ISBN 9780387970622. URL `https://books.google.com.br/books?id=w7YZAQAAIAAJ`.