CrossMark

Generating random fields with a truncated power-law variogram: A comparison of several numerical methods

Falk Heße [a,*], Vladyslav Prykhodko [a,c], Steffen Schlüter [b], Sabine Attinger [a,d]

[a] Helmholtz Centre for Environmental Research − UFZ, Computational Hydrosystems, Leipzig, Germany
[b] Helmholtz Centre for Environmental Research − UFZ, Soil Physics, Leipzig, Germany
[c] Leipzig University of Applied Sciences, Faculty of Computer Science, Mathematics and Natural Sciences, Germany
[d] University of Jena − FSU, Institute of Geoscience, Jena, Germany

ABSTRACT

In this study we describe and compare four numerical methods for the generation of random fields with a truncated power-law variogram; the Fourier method, the Randomization method, the Hybrid method as well as the Fourier−Wavelet method. We evaluate these methods with respect to their ability to represent the variogram function over a number of spatial scales as well as the Gaussianity of the generated fields. We furthermore compare these methods with respect to computational costs and investigate structural features.

Results show that the Randomization method performs well if only a few number of spatial scales (4−6 orders of magnitude) need to be represented. Due to its simpler implementation it can be preferred over the Fourier−Wavelet method. For a larger interval of spatial scales (9−12 orders of magnitude) however, the Randomization method fails to represent the variogram. Under such circumstances the Hybrid method or the Fourier−Wavelet method should be used.

The Matlab code, used for the simulations can be accessed on our institution website at http://www.ufz.de/index.php?en=32179.

1. Introduction

The main challenge for the modeling of flow and solute transport in heterogeneous media is the incorporation of the intrinsic complexity, which such a medium exhibits (Berkowitz, 2002; Neuman, 2005). Traditionally these heterogeneities were therefore conceptualized as existing only on a few, well separated, number of scales (Bear, 1972; Dagan, 1989; Rubin, 2003). This separation allows for the derivation of quasi homogeneous descriptions on the scale of interest by means of upscaling methods, like volume averaging (Whitaker, 1999) or homogenization (Hornung, 1996). Numerous studies have however shown, that such a separation of scales does not hold for many subsurface media (Boggs et al., 1992; Molz and Boman, 1995; Benson et al., 2001). A typical effect of falsely assuming such scale separation is the well known phenomena of scale-dependent effective parameters. This comprises quantities like dispersion (Neuman, 1990), permeability (Hyun et al., 2002; Raghavan, 2006) or matrix diffusion (Liu et al., 2007; Dai et al., 2007; Zhou et al., 2007). The related transport phenomenon, which arises from such heterogeneous media, is known as anomalous diffusion (Bouchaud and Georges, 1990; Metzler and Klafter, 2000). Several approaches exist, that tackle this problem (Berkowitz et al., 2006; Huang et al., 2006; Hsu and Chen, 2010; Schumer et al., 2009) with no single approach being able to fully account for it (Neuman and Tartakovsky, 2009).

Neuman (1995) showed that the problem of scale-dependent effective parameters can be attributed to the fact that many heterogeneous media exhibit fractal features. Di Federico and Neuman (1997) further demonstrated in a subsequent study that, due to the finite size of any real medium, such fractal behavior is confined by both a maximum and a minimum length scale. They showed that such problems as outlined above can be avoided by representing such media as random fields with a truncated power-law variogram. Here the power law represents the fractal nature of the medium, whereas the truncation accounts for the finite nature of real world sites. Due to the parsimonious nature and transferability of this approach, truncated power-law variograms have seen several applications in recent years (Siena et al., 2012; Guadagnini et al., 2012, 2013; Riva et al., 2013; Neuman et al., 2013).

* Corresponding author.
E-mail addresses: falk.hesse@ufz.de (F. Heße), vladyslav.prykhodko@ufz.de (V. Prykhodko), steffen.schlueter@ufz.de (S. Schlüter), sabine.attinger@ufz.de (S. Attinger).

Although the specifics of the approach of Di Federico and Neuman (1997) have been derived with respect to hydrogeology, their framework is not tied to this particular field. Any spatially varying data that (i) can be modeled by means of a Gaussian random field and (ii) shows a fractal behavior over some spatial scales, is amenable to this approach. In fact, such structures, called fractional Brownian motion (Mandelbrot and van Ness, 1968; Molz et al., 1997, 2004) as well as the related transport mechanism, i.e. anomalous diffusion, are ubiquitous in nature. One example would be the research of turbulence velocity fields (Shlesinger et al., 1987, 1993; She and LeVeque, 1994), where the often used Kolmogorov spectrum is a special case of a power law. Other fields include the topography of natural surfaces (Ouchi and Matsushita, 1992; Franceschetti et al., 1999; Lovejoy and Schertzer, 2007), soil science (Gimenez et al., 1997; Anderson et al., 1998), agronomy (He et al., 2010), wave propagation (Sahimi and Tajer, 2005; Marty and Solna, 2011) or systems biology (Caspi et al., 2000, 2002).

Due to this complex nature the numerical generation of such random fields can pose a significant challenge for random field generators and several methods have evolved over the years (Coeurjolly, 2000). An often-used class of methods is based on the decomposition of the covariance function of the random field (O'Malley et al., 2012; Cohen and Istas, 2013; Kroese and Botev, 2013). These methods are fast and reliable for smaller fields but the numerical cost tend to increase quickly for bigger realisations. Another class are so called sequential Gaussian methods (Emery, 2004) (for ready-to-use versions see for example Bellin and Rubin (1996); Deutsch and Journel (1997); Brouste et al. (2007)). These methods can easily produce Gaussian random fields, which are conditioned on point data and allow for a later refinement of the field. Major drawbacks are however, problems with correlation functions that are smooth at the origin as well as rising computational costs that arise when refining the grid. A third class are spectral methods, where a representation of the random field in the Fourier space is used for the numerical generation. Due to this spectral representation these method allow a flexible handling of specific numerical challenges, which makes this class very versatile. The most common method of these is the Fourier method (Dykaar and Kitanidis, 1992; Dietrich and Newsam, 1993; Ruan and McLaughlin, 1998; Cirpka and Nowak, 2003) but several improvements like the Randomization method (Dentz et al., 2002) have been suggested over the years. Of particular interest for the generation of fractal random field are so called hierarchical methods like the Hybrid method (Kramer et al., 2007) or Fourier–Wavelet method (Cameron, 2003).

In this study we will investigate four spectral methods for the generation of random fields: the simple Fourier method, which will serve a benchmark, the Randomization method, which is an improved version of the Fourier method as well as the Hybrid and Fourier–Wavelet methods. Due to the hierarchial nature of the last two methods we expect them to be able to reproduce a broad invterval of length scales at once, which makes them a suitable option for the generation of fractal fields. We will compare these methods with respect to their ability to reproduce the Gaussianity as well as the variogram functions used for the generation of the random fields. In addition to these analytical measures we also investigate structural properties of the random fields. Since no theoretical predictions of these properties are known we check these measures with respect to plausibility.

The rest of this paper is organized as follows; In Section 2 we describe the methods used in this study. This comprises a short derivation of the truncated power-law variogram and its associated spectra. Then we derive and describe the numerical methods used to generate random fields. They are as follows; the Fourier method, the Randomization method, the Hybrid method as well as the Fourier–Wavelet method. In addition we detail the different measures, which were used in order to assess the quality of the numerically generated random fields. In Section 3 we present and discuss the results of our findings. This includes a comparison of the computational costs associated with each method, the quality of the generated random fields as well as an investigation on their structural properties. In the last section we give a summary of our main findings and provide a short outlook.

2. Methodology

Due to the intrinsic complexity of heterogeneous media as well as the uncertainty of field studies it is common to use random fields, i.e. spatially-dependent random variables, in order to describe certain geological properties (for a primer into the notation used herein we refer to Appendix A). Such random fields are commonly described in geostatistics by their expectation value and their variogram, the latter of which will be detailed in the following.

2.1. Variograms and associated spectra of random fields

In geostatistics the variogram function for a specific geologic material is usually acquired by estimation from a sample according to Equation (A.2) and subsequent fitting to a conceptual model function. Several model, which are relevant to this study are explained in the following.

2.1.1. Classic variograms and spectra

The variogram functions, which are used most often in geostatistics are the exponential and the Gaussian variogram. The variogram and spectrum of the former are defined as

$$\gamma(h, \lambda) = \sigma^2 \left[1 - e^{-\frac{|h|}{\lambda}} \right], \tag{1a}$$

$$S(\mathbf{k}, \lambda) = \sigma^2 \lambda^d \frac{\Gamma\left[\frac{d+1}{2}\right]}{\left(\pi\left(1 + \mathbf{k}^2\lambda^2\right)\right)^{\frac{d+1}{2}}} \tag{1b}$$

as well as for the latter

$$\gamma(h, \lambda) = \sigma^2 \left[1 - e^{-\frac{\pi}{4}\left(\frac{h}{\lambda}\right)^2} \right], \tag{2a}$$

$$S(\mathbf{k}, \lambda) = \sigma^2 \left(\frac{\lambda}{\pi}\right)^d e^{-\frac{1}{\pi}(\mathbf{k}\lambda)^2}, \tag{2b}$$

with d being the spatial dimension ($d = 1,2,3$). Both variogram functions assume a single characteristic length scale h λ on which the observed heterogeneities appear. A condition, which is not met by many porous and fractured materials (Ritzi et al., 2004). Neuman et al. (2008) argue that the apparent ability of these models to characterize a wide range of porous materials can be seen as an artifact of the finite nature of the sampling process itself.

2.1.2. Power-law variogram and spectrum

Exponential and Gaussian variograms assume the existence of a single length scale on which heterogeneities of the geological medium appear. It has however, been shown that many effective quantities like permeability or dispersivity of geologic media depend on the length scale on which they are derived (Neuman, 1994; SanchezVila et al., 1996; Schulze-Makuch et al., 1999). Plotting these quantities vs. the scale of observation they often show a linear increase in a logarithmic scale. Such a linear behavior is consistent with the assumption of a power law for the variogram and the spectrum of the underlying medium

$$\gamma(h) = C_0 |h|^{2H} \quad 0 < H < 1, \tag{3a}$$

$$S(\mathbf{k}) = \frac{\Gamma(1 + 2H)\sin(\pi H)}{\pi} C_0 |\mathbf{k}|^{-(1+2H)} \quad 0 < H < 1. \tag{3b}$$

Here C_0 is a constant and H is the Hurst coefficient (Hurst, 1951). An often used power-law spectrum is the Kolmogorov spectrum having the following variogram and spectral density function

$$\gamma(h) = -2C_0 \Gamma\left(-\frac{2}{3}\right) |2\pi h|^{\frac{2}{3}} \approx 8.0368 C_0 |2\pi h|^{\frac{2}{3}}, \tag{4a}$$

$$S(\mathbf{k}) = C_0 |\mathbf{k}|^{-\frac{5}{3}}. \tag{4b}$$

According to Equation (3a) we get a Hurst coefficient of $H = 1/3$. The Kolmogorov spectrum was used in this study as a benchmark due to its wide use (Hunt et al., 1991; Cameron, 2003; Kramer et al., 2007).

Di Federico and Neuman (1997) could show that a generic power law as given by Equations (3a) and (3b) can be decomposed into an integral representation with weighted modes consisting of either Gaussian or exponential functions

$$\gamma(h) = C \int_0^\infty \frac{\gamma(h,n)}{n^{1+2H}} dn, \tag{5a}$$

$$S(\mathbf{k}) = C \int_0^\infty \frac{S(\mathbf{k},n)}{n^{1+2H}} dn. \tag{5b}$$

Here the parameter $n = 1/\lambda$ is the wave number, i.e. the inverse of the correlation length λ of each mode, and C is a constant with the dimensions $[L^{-1-2H}]$. The power law resulting from Equation (5a) will then follow Equation (3a) with

$$C_0 = C \frac{\Gamma(1-2H)}{2H}, \quad 0 < H < 1/2, \tag{6}$$

for the case of exponential modes and

$$C_0 = C \frac{\Gamma(1-H)}{2H} \left(\frac{\pi}{4}\right)^H, \quad 0 < H < 1, \tag{7}$$

for the case of Gaussian modes respectively (in both cases the variance of the single modes was assumed to be unity $\sigma^2 = 1$).

This decomposition can be interpreted such that fractal geological media is composed of an infinite hierarchy of scales, each of which represented by a classic variogram (Neuman and Di Federico, 2003). The major flaw of this approach, or any fractal model for that matter, is the fact that real media has always both an upper λ_u as well as a lower length scale λ_l confining the fractal behavior within an interval of scales. Such a behavior can however, be represented by truncating the integral in Equation (5a) resulting in a truncated power-law variogram.

2.1.3. Truncated power-law variograms and spectra

As shown by Di Federico and Neuman (1997) the impact of the upper length scale λ_u can be represented by the truncation of the integral from Equation (5a) at a lower wave number $n_l = 1/\lambda_u$. The resulting (single) truncated power-law variograms and spectra are then

$$\gamma_{n_l}(h) = C \int_{n_l}^\infty \frac{\gamma(h,n)}{n^{1+2H}} dn, \tag{8a}$$

$$S_{n_l}(\mathbf{k}) = C \int_{n_l}^\infty \frac{S(\mathbf{k},n)}{n^{1+2H}} dn. \tag{8b}$$

Unlike in case of Equation (5a), where both exponential and Gaussian modes yielded the same expression, the resulting truncated power law is different for both types.

Exponential modes. In case of exponential modes, i.e. Equations (1a) and (1b) are inserted into Equations (8a) and (8b), the following variogram and spectrum result

$$\gamma_{n_l}(h) = \sigma^2 \left[1 - e^{-|h|n_l} + (|h|n_l)^{2H} \Gamma(1-2H, |h|n_l)\right], \tag{9a}$$

$$S_{n_l}(\mathbf{k}) = \sigma^2 \frac{2H\Gamma\left[\frac{d+1}{2}\right]}{\pi^{\frac{d+1}{2}}(d+2H)n_l^d} {}_2F_1\left[\begin{matrix}\frac{d+1}{2}, \frac{d+2H}{2} \\ \frac{d+2+2H}{2}\end{matrix} ; -\left(\frac{|\mathbf{k}|}{n_l}\right)^2\right]. \tag{9b}$$

Here $\Gamma(\cdot, \cdot)$ is the incomplete gamma function, ${}_2F_1$ is known as Gauss's hypergeometric function and the variance σ^2 is given as

$$\sigma^2 = \frac{C}{n_l^{2H} 2H}. \tag{10}$$

Gaussian modes. In case of Gaussian modes, i.e. Equations (2a) and (2b) are inserted into Equations (8a) and (8b), the following variogram and spectrum result

$$\gamma_{n_l}(h) = \sigma^2 \left[1 - e^{-\frac{\pi}{4}(hn_l)^2} + \left(\frac{\pi}{4}(hn_l)^2\right)^H \Gamma\left(1-H, \frac{\pi}{4}(hn_l)^2\right)\right], \tag{11a}$$

$$S_{n_l}(\mathbf{k}) = \sigma^2 \frac{Hn_l^{2H}\pi^{-\frac{d}{2}+H}}{|\mathbf{k}|^{2\left(\frac{d}{2}+H\right)}} \left(\Gamma\left[\frac{d+2H}{2}\right] - \Gamma\left[\frac{d+2H}{2}, \frac{1}{\pi}\left(\frac{|\mathbf{k}|}{n_l}\right)^2\right]\right). \tag{11b}$$

Comparing Equations (9a) and (11a) shows a very similar behavior for both expressions. Both exhibit a power-law term, which is important for small lengths ($h < \lambda_u$). In this region both variograms are similar to a power law (Equation (3a)), which is exemplified by a linear increase in a log–log plot having a slope of $2H$ (see
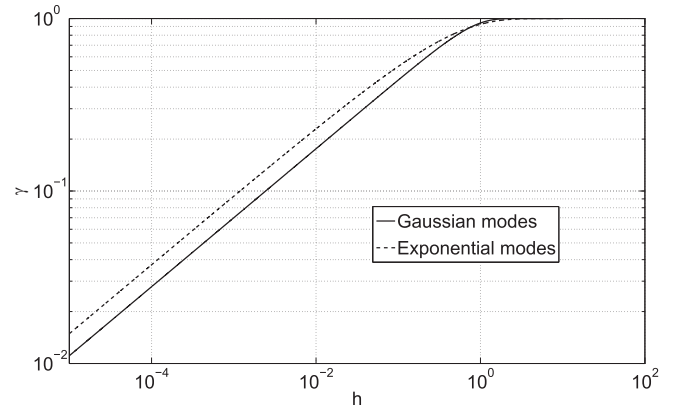

Fig. 1. Example truncated power-law variograms with Gaussian (solid line) and exponential modes (dashed line). The parameters were: $\lambda_u = 1$ and $H = 0.25$.

Fig. 1). Both expressions also exhibit an exponential and a Gaussian term, respectively. As a result the variograms are saturating for bigger lengths ($h > \lambda_u$). The incomplete gamma function is mediating the transition between these two regions.

Despite similar variograms, the spectra of both models are very different (compare Equations (9b) and (11b)).

Double Truncation. In the analysis above we considered the impact of an upper cut-off length λ_u representing the finite size of real geological media. In order to complete the description the lower cut-off length λ_l, representing the finite resolution, has to be considered as well. The variograms and associated spectra are given by

$$\gamma_{n_l, n_u}(h) = C \int_{n_l}^{n_u} \frac{\gamma(h,n)}{n^{1+2H}} dn = \gamma_{n_l}(h) - \gamma_{n_u}(h) \tag{12}$$

$$S_{n_l, n_u}(\mathbf{k}) = C \int_{n_l}^{n_u} \frac{S(\mathbf{k},n)}{n^{1+2H}} dn = S_{n_l}(\mathbf{k}) - S_{n_u}(\mathbf{k}), \tag{13}$$

with the variance given as

$$\sigma^2_{n_l, n_u} = \sigma^2_{n_l} - \sigma^2_{n_u} = \frac{C}{\left(n_l^{2H} - n_u^{2H}\right)2H}. \tag{14}$$

Equations (12)–(14) are therefore completing the description for the truncated power-law variograms and spectra.

Note that in the following we will use the single-truncated power law exclusively. As a result we will the use the general expression truncated power law.

2.2. Numerical generation of random fields

In this study four different numerical methods were used for the generation of Gaussian random fields; the Fourier method, the Randomization method, the Hybrid method and the Fourier–Wavelet method. Due to computational costs one-dimensional fields, only, were used for the investigation of the variogram structure and kurtosis, whereas for the investigation of the structural properties a smaller set of two-dimensional random fields was used. The four aforementioned methods are therefore introduced in the following for one-dimensional fields with an extension to two- and three-dimensional media being given in Appendix C.

A common feature of these four methods is the fact, that they are spectral methods, i.e. they are based on a spectral representation of a random field (see Appendix B). Within such a context we can indentify two challenges: (i) the quadrature of the integral of the spectrum and (ii) the numerical representation of white noise. How these challenges are meet by the different methods is detailed in the following.

2.2.1. The Fourier method

The Fourier method is the simplest and most robust discretization scheme, which was considered in this study. It can be understood as a Riemann sum discretization of Equation (B.3) and is therefore easy to implement. It does however, suffer from several flaws (Elliott et al., 1997), which will be highlighted in the following. The basic idea behind the method is the truncation (not to be confused with the truncation mentioned above) and discretization of the integral according to the mid-point rule with equi-spaced grid points. Using for the truncation a finite symmetric partition of N intervals having equal widths Δk and inserting this into Equation (B.3) results in the following approximation
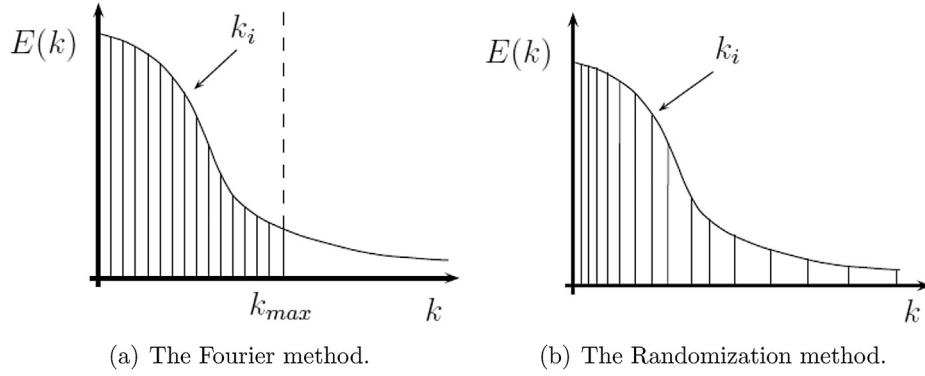
Fig. 2. Schematic of the sampling procedure in the different methods. (a) The Fourier method. (b) The Randomization method.

$$u(x) \approx \sqrt{2\sigma^2} \sum_{i=1}^{N} \sqrt{E(k_i)} \left(\cos(2\pi k_i x)\Delta W_i^1 + \sin(2\pi k_i x)\Delta W_i^2 \right). \tag{15}$$

The calculation of Equation (15) is straight forward except for the determination of the discretized white noise

$$\Delta W_i^n = Z_i^n \sqrt{\Delta k}, \tag{16}$$

with Z_i^n being independent normal random variables. Introducing this relationship into of Equation (15) results in the representation of Equation (B.3) according to the Fourier method

$$u(x)_{\text{four}} = \sqrt{2\sigma^2} \sum_{i=1}^{N} \sqrt{E(k_i)} \left(Z_i^1 \cos(2\pi k_i x) + Z_i^2 \sin(2\pi k_i x) \right) \sqrt{\Delta k}. \tag{17}$$

One of the aforementioned drawbacks of the Fourier method is the artificially introduced periodicity with length $(\Delta k)^{-1}$. This artifact is due to the finite resolution of the scheme in the Fourier domain and cannot be directly avoided.

Another problem is caused by the truncation of the integral of Equation (B.3) assuming the integrand to be zero (or almost zero) beyond this truncation. A condition, which is met by the classic spectra given in Section 2.1.1 but not necessarily by those given in Sections 2.1.2 and 2.1.3.

2.2.2. The Randomization method

The Randomization method has been developed by Sabelfeld and his coworkers (Kramer et al., 2007). It can be understood as a modification of the aforementioned Fourier method, whereby the equi-distant approximation of the stochastic integral in Equation (B.1) is replaced by a random discretization (henceforth the name). Using Equation (B.3) both terms can be treated separately. The first term is then

$$u^1(x) = \sqrt{2\sigma^2} \int_0^\infty \cos(2\pi k x) \sqrt{E(k)} \, dW^1(k), \tag{18}$$

which is again a random field with zero mean and the variance

$$\text{Var}\left(u^1(x) \right) = \left\langle \left| u^1(x) \right|^2 \right\rangle = 2\sigma^2 \int_0^\infty \cos^2(2\pi k x) E(k) \, dk. \tag{19}$$

This integral can be approximated by means of a Monte-Carlo integration resulting in the following sum

$$\text{Var}\left(u^1(x) \right) \approx 2 \frac{\sigma^2}{N} \sum_{i=1}^{N} \cos^2(2\pi k_i x). \tag{20}$$

Here the k_i are random numbers following the density distribution function defined by $E(k)$. By virtue of this relationship Equation (18) can be approximated according to

$$u^1(x) \approx \sqrt{2 \frac{\sigma^2}{N}} \sum_{i=1}^{N} Z_i^1 \cos(2\pi k_i x), \tag{21}$$

where the Z_i^1 are again independent Gaussian random variables. Applying the same procedure to the second term in Equation (B.3) results in the Randomization method

$$u(x)_{\text{rand}} = \sqrt{2 \frac{\sigma^2}{N}} \sum_{i=1}^{N} \left(Z_i^1 \cos(2\pi k_i x) + Z_i^2 \sin(2\pi k_i x) \right). \tag{22}$$
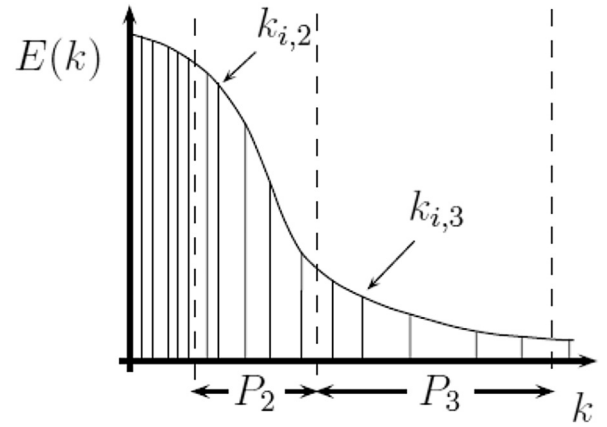

Fig. 3. Schematic of the sampling procedure in the Hybrid method.

Equation (22) is structurally very similar to Equation (17). The main difference is the way the k_i are determined.

In the Fourier method these equi-distant points k_i are sampling the finite interval $[0, k_{max}]$ with N partitions having the width Δk (see schematic in Fig. 2a). For a spectrum with long tails such a procedure can be disadvantageous since k_{max} must be very big to sample a sufficiently large portion of the spectrum. By randomly choosing the sampling points k_i the Randomization method circumvents this problem (see schematic in Fig. 2b). As a result the interval being sampled is not confined by a finite k_{max} but is (potentially) the whole support of the spectral density function $E(k)$ (for the definitions given above this is the whole domain $[0, \infty)$). Higher regions of the Fourier domain are therefore less likely to be sampled but not discarded completely as in case of the Fourier method.

2.2.3. The Hybrid method

The Hybrid method can be understood as a combination of the concepts of the Fourier method (deterministic sampling of the Fourier domain) and the Randomization method (random sampling of the Fourier domain) by applying a stratified sampling procedure to the latter. To that end the Fourier domain is partitioned into N_P non-empty and non-overlapping parts $\bigcup_{j=1}^{N_P} P_j = [0, \infty)$. Within each part P_j, having the boundary values $P_j = [k_{P_j}, k_{P_{j+1}})$, a random sampling of N points $k_{i,j}$ is applied according to the Randomization method. The PDF within every part P_j naturally has to be adapted. It now reads

$$E_j(k) = \begin{cases} \frac{2}{\sigma_j^2} E(k), & \text{for } k \in P_j, \\ 0, & \text{else}, \end{cases} \tag{23}$$

where the variance σ_j^2 of each part is computed according to

$$\sigma_j^2 = 2 \int_{k_{P_j}}^{k_{P_{j+1}}} E(k) \, dk. \tag{24}$$

Inserting this partitioning into Equation (22) results in the following formulation

$$u(x)_{\text{hyb}} = \sum_{j=1}^{N_P} \sqrt{2\frac{\sigma_j^2}{N}} \sum_{i=1}^{N} \left(Z_{i,j}^1 \cos(2\pi k_{i,j} x) + Z_{i,j}^2 \sin(2\pi k_{i,j} x) \right). \tag{25}$$

Here the random numbers $Z_{i,j}^n$ are again independent standard Gaussian random variables.

The main advantage of this stratified sampling technique is the possibility of sampling specific portions of the Fourier domain (usually those with high values of k) with higher accuracy. This further improves the advantages of the Randomization method. The Fourier method had the drawback of a truncation of the stochastic integral for high values of k (see schematic in Fig. 2a). As explained this can prove to be problematic for spectra having long tails. The Randomization method could circumvent this problem since values from the whole Fourier domain would be sampled according to the PDF given by the respective spectral density function (see schematic in Fig. 2b). This procedure is a clear improvement but still leaves regions with small values of $E(k)$ with smaller sample size and therefore smaller accuracy. The Hybrid method is countering this tendency by sampling each part P_j with the same sample size (see schematic in Fig. 3). The combination of both methods (Fourier method and Randomization method) improves the representativeness and therefore the overall quality of the sampling procedure.

The crucial question for the application of the Hybrid method is the choice of the partitioning procedure. According to a variety of performed test runs (not discussed here) and in accordance with the results of Kramer et al. (2007) we found that a logarithmic partitioning, with doubling interval lengths, gave best results. We therefore used $P_1 = [0,1)$ for the first part and $k_{P_{j+1}} = 2k_{P_j}$ for every successive interval boundary. Please note that Kramer et al. (2007) call this specific partitioning strategy for the Hybrid method the 'variant C of the Randomization method' in their paper.

2.2.4. The Fourier–Wavelet method

The last method, which was investigated in our study, is the Fourier–Wavelet method, which has been developed by Elliott et al. (1997). As the name implies wavelets are applied for the discretization of Equation (B.1) in particular for the random measure $dW(k)$. In contrast to Equation (16) the following ansatz is used for the representation of the increments of the Wiener process

$$dW(k) = \sum_i Z_i \phi_i(k) dk. \tag{26}$$

Here Z_i are again independent Gaussian random variables and the ϕ_i must form an orthonormal basis. The main idea of Elliott et al. (1997) is to use Fourier transformed Wavelets ψ_i or more specifically Meyer Wavelets for $\phi_i = \mathscr{F}[\psi_i]$ (henceforth the name of the method). The Fourier transform of the Meyer wavelet has the form

$$\Phi(k) = \mathscr{F}[\psi](k) = -\mathrm{i}\ \text{sign}(k) e^{i\pi k} b(|k|), \tag{27}$$

with $b(k)$ defined as

$$b(k) = \begin{cases} \sin(\frac{\pi}{2}\nu(3k-1)), & \text{if } k \in (\frac{1}{3}, \frac{2}{3}], \\ \cos(\frac{\pi}{2}\nu(3k/2-1)), & \text{if } k \in (\frac{2}{3}, \frac{4}{3}], \\ 0, & \text{else.} \end{cases} \tag{28}$$

The function $\nu(k)$ must be a nondecreasing smooth function with the following properties

$$\nu(k) = \begin{cases} 0, & \text{if } k < 0, \\ 1, & \text{if } k > 1, \\ \nu(k) + \nu(1-k) = 1, & \text{else.} \end{cases} \tag{29}$$

Elliott et al. (1997) or Cameron (2003) propose the use of Spline functions with such properties.

We generate the family of wavelet functions $\psi_{i,j}(x)$ from the Mother Wavelet $\Psi(x)$ by the following wavelet relation

$$\psi_{i,j}(x) = 2^{\frac{i}{2}} \psi\left(2^i x - j\right), \text{with } i, j = 0, \pm 1, \pm 2, \cdots \tag{30}$$

using translations and stretchings. This set of wavelets is forming an orthonormal basis, a property, which is preserved under the Fourier transform. We have therefore found a set of functions applicable for Equation (26)

$$dW(k) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} Z_{i,j} \phi_{i,j}(k) dk. \tag{31}$$

Inserting Equation (31) into Equation (B.1) and re-arranging results in the following representation of a random field

$$u(x) = \sqrt{2\sigma^2} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} Z_{i,j} \int_{-\infty}^{\infty} e^{-2\pi kx} \sqrt{E(k)} \phi_{i,j}(k) dk. \tag{32}$$

Note that Equation (32) is *not* an approximation of Equation (B.1) but a transformation into a different basis. In order to be used as a numerical algorithm practical truncations must be introduced for both sums as well as a method for the calculation of the Fourier integral.

In order to find an appropriate truncation for the index i the length scales of $u(x)$, which need to be represented, must be considered. With respect to the single-truncated power law given by Equations (9a) and (11a) as well as the double-truncated power law given by Equation (12) we can identify this length scales with the upper λ_u and lower truncation length λ_l. By virtue of truncating N the Fourier–Wavelet method provides a very precise handle for the exact representation of the desired length-scales. For the truncation of the index j the finite support of the wavelets can be used. For the determination of a single value of the random field $u(x)$ only $2M$ terms need to be considered, for which $\left|j - 2^i x\right| < M$ holds. As a result the Fourier–Wavelet method is described by the following finite double sum

$$u(x)_{fw} = \sqrt{2\sigma^2} \sum_{i=0}^{N} \sum_{j=1-M}^{M} Z_{i,j} c_{i,j}(x), \tag{33}$$

with $c_{i,j}(x) = \mathscr{F}^{-1}[\sqrt{E}\Phi_{i,j}](x)$. In order to complete the derivation of the Fourier Wavelet method the coefficient functions $c_i(x)$, given by inverse Fourier transform of $\sqrt{E}\Phi_i$, need to be specified.

2.3. Assessing the quality of numerically generated random fields

In order to compare the different methods for the numerical generation of random fields we investigated the quality of the generated fields with respect to several criteria. The criterion for the one-point distribution, i.e. the Gaussianity of every point of $u(\mathbf{x})$, was the kurtosis, whereas the two-point distribution was investigated by comparing the estimating variogram with the analytical function. Both of these criteria were assessed with respect to a wide interval of decades.

2.3.1. Estimating the kurtosis

The kurtosis G_4 is commonly described as a measure for the peakedness or the slope of the PDF of the random variable. It can be used for assessing the Gaussianity of the one-point distribution of a random field. The kurtosis is defined as the fourth centralized moment μ_4 of the random variable X normalized by the square of the variance

$$G_4(X) = \frac{\mu_4(X)}{\sigma^4(X)}. \tag{34}$$

The kurtosis of a (one-dimensional) random field $u(x)$ is defined

$$G_4(X) = \frac{E\left[(u(x) - u(0))^4\right]}{E\left[(u(x) - u(0))^2\right]^2}. \tag{35}$$

A simple method for the estimation of the kurtosis for a given sample of N realizations $u_n(x)$ is to estimate the second and fourth moment and using the ratio

$$\overline{G}_4(x) = N \frac{\sum_{n=1}^{N} u_n^4(x)}{\left(\sum_{n=1}^{N} u_n^2(x)\right)^2}. \tag{36}$$

Since the kurtosis of a Gaussian random variable is always $G_4(X) = 3$ this estimate was used as a measure of the Gaussianity of the generated random fields.

2.3.2. Estimating the variogram

For the assessment of the two-point distribution of the generated random fields we used the variogram as defined in Equation (A.2). For a given sample of N realizations $u_n(x)$ we estimated the variogram by the following estimator

$$\overline{\gamma}(h) = \frac{1}{N} \sum_{n=1}^{N} |u_n(x+h) - u_n(x)|^2. \tag{37}$$

2.3.3. Estimation of reproduced decades

The generated numerical fields were supposed to exhibit truncated power-law structures, i.e, both the kurtosis and variogram should have been reproduced over a wide range of scales (see Fig. 1). In order to quantify this statement we used the relative error between the expected $\gamma(h)$ and the estimated variogram $\overline{\gamma}(h)$

$$\varepsilon = \frac{|\gamma(h) - \overline{\gamma}(h)|}{\gamma(h)}. \tag{38}$$

In order to guarantee that $\overline{\gamma}(h)$ is close to the real variogram of the generated random fields we used very high sampling sizes (several thousands of realizations). For the derivation of the reproduced number of decades we had to determine the set h_ε for which the relative error was smaller than a maximally acceptable threshold

$\varepsilon_{\max}$. This threshold was set to 0.1, i.e. a relative error of 10%. The set of points for which the estimated variogram is acceptable was therefore defined as

$$h_\varepsilon = \{h|\varepsilon(h) < \varepsilon_{\max}\}. \tag{39}$$

In order to determine a contiguous interval Δh with upper $h_{\max}$ and lower $h_{\min}$ boundaries we used $h_{\max} = \max_h h_\varepsilon$ as the upper boundary. For the lower boundary we defined $h_{\min}$ such that a certain percentage p_h of Δh belongs to h_ε. This percentage value was set to $p_h = 90\%$. With these boundaries we could compute the number of reproduced decades d_ε according to

$$d_\varepsilon = \log_{10} \frac{h_{\max}}{h_{\min}}. \tag{40}$$

It should be noted at this point that under regular circumstances the number of decades, that need to be represented in a numerical simulation of flow and transport in heterogeneous media is rather confined. For a simple finite-difference discretization a number of four decades will results in 10^8 grid points in two and 10^{12} in three dimension. However, with the development of newer techniques this number can be much higher. Such techniques include adaptive grid refinement (Trangenstein, 2002) or multi-scale codes (Juanes and Tchelepi, 2008; Nordbotten, 2009).

2.3.4. Analyzing structural properties

In addition, to the investigation of criteria that are known a priori, e.g. kurtosis and variogram, random fields can also be judged by morphological properties of the generated random fields. In this study we focused on possible failures of the generation methods to reproduce the shape of clusters in the fractal regime.

To that end we used two-dimensional realizations on different spatial scales, i.e. different domain sizes L_{num}, whereas the upper truncation length remained constant at $\lambda_u = 1$ (for an overview on the generation of higher-dimensional random fields we refer to Appendix C). These realizations were thresholded into level sets, i.e. binary representations, where areas with values higher than the threshold are referred to as clusters or objects and the rest of the domain as background, respectively. To make the shape properties location-independent we adapted the threshold for each realization to its mean gray value.

Three arbitrary results of this procedure can be seen in Fig. 4, where every example had a different domain size L_{num}. The examples were two-dimensional random fields with a truncated power law having Gaussian modes. In Fig. 4a the size of the domain L_{num} was smaller than the cut-off length λ_u, i.e. the field was purely fractal. This can be seen by the fact that clusters of low or high permeability are penetrating the whole domain. Zooming into the original field at any given point would consequently reveal new similar structures. In Fig. 4b L_{num} was bigger than λ_u and the random field looks therefore structurally different. The clusters of either low or high permeability are now finite. Zooming out to 4(c) eventually lead to structures indistinguishable from a classic random field with a Gaussian variogram function.

These random fields were subsequently characterized with respect to three different criteria: (i) the thickness of clusters, represented by the chord length, (ii) the roughness of clusters, represented by the boundary length as well as (iii) its connectivity.

Thickness (chord length). A plausible estimator for the thickness of clusters is the chord length distribution. A chord is a straight line that connects two opposing boundaries of an object. A chord length distribution $p_c(h)$ is defined as the probability that a randomly chosen chord in a given phase has a length in the range of $[\mathbf{x},\mathbf{x}+h]$, where $\mathbf{x}$ is the location vector and h is distance (Roberts and Torquato, 1999; Torquato, 2002; Schlueter and Vogel, 2011). The chord length distribution of the white phase displays the thickness of objects, whereas the chord length distribution of the black phase renders the separation between objects (see Fig. 4). A mean chord length is defined by integrating these frequency distributions over all chord lengths,

$$\langle p_c \rangle = \int p_c(h) \mathrm{d}h. \tag{41}$$

Roughness (boundary length). In addition to the mean chord length we also investigated the roughness of clusters, i.e. the boundary length L_B between objects and background. Boundary length represents one of the Minkowski functionals defined on binary sets in two-dimensional space (Mantz et al., 2008; Mecke, 2000; Vogel et al., 2010). In contrast to counting pixel edges, employing Minkowski functionals guarantees an unbiased estimate of boundary lengths in spite of the pixilized nature of the domain.

Connectivity. Amongst the various metrics that describe the connectivity of a binary set (Renard and Allard, 2011), two were used in our study.

The first connectivity measure that we used, was the two-point cluster function (Torquato et al., 1988; Jiao et al., 2009), sometimes also simply denoted as (pair) connectivity function (Allard, 1993; Western et al., 2001; Schlueter and Vogel, 2011; Renard and Allard, 2011). It is defined as the distance-dependent probability that two points are located within the same cluster, i.e. a path of arbitrary shape between the two points $[\mathbf{x},\mathbf{x} + h]$ exists (denoted as $\mathbf{x} \leftrightarrow \mathbf{x} + h$) that lies entirely within a subset of phase α in the domain Ω,

$$\tau(h) = \mathrm{Prob}\{\mathbf{x} \leftrightarrow \mathbf{x} + h | \mathbf{x} \in \alpha, \mathbf{x} + h \in \Omega\}. \tag{42}$$

Note that pixels are considered to be connected by both faces and corners, i.e. each pixel has eight connected neighbors. For non-percolating sets this probability reduces to zero as soon as the distance between the two points h exceeds the size of the largest cluster. In this case a mean connectivity range can be defined as the integral scale of the pair connectivity function $\langle \tau \rangle = \int \tau(h) \mathrm{d}h$ similar to the mean chord length above. For percolating sets the probability of a connection over large distances converges to the area fraction of the largest cluster and the mean connectivity length merely becomes an artifact of the finite domain size. Thus, $\langle \tau \rangle$ diverges drastically around the percolation threshold.

The second connectivity measure that we used, was the proportion of pairs of cells, that are connected amongst all the pairs of cells within a phase (Renard and Allard, 2011),

$$\Gamma = \frac{1}{n_\alpha^2} \sum_{i=1}^{N_\alpha} n_i^2. \tag{43}$$

Here n_α is the total number of cells in phase α and n_i is the number of cells within a specific cluster and N_α is the total number of clusters. Hence, the second-order central moment of the cluster size distribution was considered in our study. Γ is small for many disconnected clusters of similar size and high if a large percolating cluster exists.

3. Results and discussion

In this section the results are presented, which have been obtained by testing and comparing the Fourier method, the Randomization method, the Hybrid Method and the Fourier—Wavelet method with respect to several criteria. This comprises computational costs as well as the quality of the generated fields with respect to several statical characteristics.

3.1. Computational costs

Computational costs are a very important criterion for the classification of numerical methods. The quality of a generated
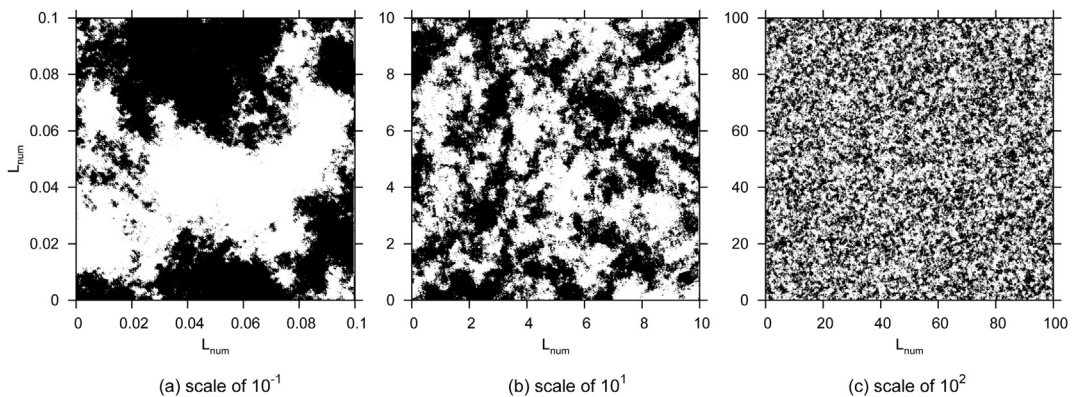

Fig. 4. Level set for one realization with Gaussian modes and L_{num} of 10^0 (left), 10^1 (middle) and 10^2 (right) thresholded at the mean gray value of the random field.

random field for a specific method is generally a function of the numerical workload. In order to decide, which method is most favorable for a desired quality the computational cost has therefore to be taken into account. It is furthermore clear, that the actual running time for generation of random fields may differ dramatically for a given setting, depending on the hardware, the programming language used for the code and the implementation details of the method itself. As a result only the number of operations, which are needed for every method in order to generate a single random value u at a single given point x, were determined. This number depended on the number of terms of the sum, which were used to approximate the stochastic integral given by Equation (17).

3.1.1. The Fourier method

According to Equation (17) the Fourier method is approximating the stochastic integral by a single finite Riemann sum. The relevant parameter was therefore the number of the terms of this sum N. For every term two independent Gaussian random variables had to be generated and two trigonometric functions needed to be evaluated. In addition, we had to evaluate the spectrum $S(k)$ for every term. Depending on the implementation these costs could be smaller or much bigger compared to the other numerical computations. Owing to this fact this issue is discussed in a dedicated section below.

3.1.2. The Randomization method

In analogy to the Fourier method the stochastic integral is approximated by a finite sum, where in every term we had to generate two independent Gaussian random variables as well as to evaluate two trigonometric functions. The grid points k_i in every terms however, were no longer chosen deterministically but were random variables themselves. The PDF of these k_i is given by the spectral density function $E(k)$. Depending on the method, which was chosen for the generation of the k_i this part could distinctly increase the computational burden compared to the Fourier method. In our case of we chose *Inverse Transform Sampling*, i.e. we had to evaluate the cumulative distribution function of $E(k)$ several times. The exact number of evaluations depended on the convergence of the method as well as the desired accuracy. Depending on the used variogram function the evaluation of the associated CDF could be very costly (see below). In these cases the Randomization method was much slower than the Fourier method.

3.1.3. The Hybrid method

The Hybrid method is structurally very similar to the Randomization method (compare Equation (22) to Equation (25)). As a result above statements about the Randomization method can be transferred to the Hybrid method, too. The number of terms that had to be computed was however, $N\,N_P$. The PDF had to be computed according to Equations (23) and (24), which in turn did increase the number of operations and thus computational costs.

3.1.4. The Fourier−Wavelet method

The Fourier−Wavelet method is significantly more complicated compared the other three investigated methods. The stochastic Fourier integral given be Equation (B.1) was approximated by two finite sums with independent Gaussian random numbers as weights. The kernel functions ware evaluated at equi-distant grid points and could be used for all further calculations. This procedure could be performed as pre-processing. The generation of the random weights was more difficult compared to the other methods, since their computation depended on the grid point x. In total $2N(M+1)$ random numbers were needed for the generation of a single value of u at x.

3.1.5. Computation of the spectra and associated CDF of the truncated power laws

As explained above we needed to evaluate the spectrum $E(k)$ in case of the Fourier and Fourier−Wavelet method as well as the $CDF(k)$ in case of the Randomization and Hybrid method. Using Equations (11b) and (9b) we could identify the incomplete Gamma function as well as Gauss's hypergeometric function as the only non-trivial mathematical expressions, which needed to be evaluated. Most modern computational environments provide fast and reliable implementations of the incomplete Gamma function. This is in contrast to Gauss's hypergeometric function (or other hypergeometric functions for that matter). Tests with ready-to-use implementations in Matlab and Mathematica showed much higher computational costs compared to the incomplete Gamma function as well as problems with stability. The generation of a random field with an truncated power law with exponential modes took therefore much longer than a field with Gaussian modes or did even fail completely some times.

As a result we investigated an alternative by using approximations of Gauss's hypergeometric function. Several approaches exist like numerically solving the associated ODE (Pearson, 2009), a numerical quadrature of the associated integral (Gautschi, 2002), recurrence relations (Gil et al., 2006; Ibrahim and Rakha, 2008), Padé approximations amongst others (for an overview see Pearson (2009)). Our tests showed that for the parameter and variable regime, which prevailed in our scenarios, the Laplace approximation yielded the most suiting results (results not shown). This method, developed by Butler and Wood (2002), approximates the integral representation of Gauss's hypergeometric function

$$
{}_2F_1\begin{bmatrix} a & b \\ & c \end{bmatrix}; x = B(a, c-a)^{-1} \int_0^1 t^{a-1}(1-t)^{c-a-1}(1-xt)^{-b}\mathrm{d}t,
$$

(44)

with $B(d,e)$ being the beta function ((Abramowitz and Stegun, 1970)). The integral from Equation (44) can be brought into the following form

$$
I = \int h(t)e^{-\lambda g(t)}\mathrm{d}t,
$$

(45)

where λ is a real-valued parameter. This integral is known to have the following approximation in one dimension

$$
\tilde{I} = \sqrt{\frac{2\pi}{\lambda\left|\frac{\mathrm{d}^2}{\mathrm{d}t^2}g(t_{\min})\right|}}h(t_{\min})e^{-\lambda g(t_{\min})}.
$$

(46)

Here $t_{\min}$ is the minimum of y within the interval $(0,1)$. With respect to the coefficients used in Equation (44) this value can be computed according to the following expression

$$
t_{\min} = \frac{2a}{\sqrt{(x(b-a)-c)^2 - 4ax(c-b)} - (x(b-a)-c)}.
$$

(47)

With this information we could approximate Gauss's hypergeometric function with the following expressions

$$
{}_2F_1\begin{bmatrix} a & b \\ & c \end{bmatrix}; x = \frac{{}_2F_1(a,b,c;x)}{{}_2F_1(a,b,c;0)},
$$

(48a)

$$
{}_2F_1(a,b,c;x) = \sqrt{\frac{2\pi}{j}}\,\frac{t_{\min}^a}{B(a,c-a)}\frac{(1-t_{\min})^{c-a}}{(1-xt_{\min})^b},
$$

(48b)

$$j = a(1 - t_{\min})^2 + (c - a)t_{\min}^2 + bx^2 \frac{(1 - t_{\min})^2}{(1 - xt_{\min})^2}. \qquad (48c)$$

Equations (48a), (48b) and (48c) only contain algebraic expressions. The approximation was therefore easy to implement, fast as well as stable.

The quality of the approximation was very good in all investigated cases, with a relative error of a few per mille at the most (see Fig. 5 a). We also investigated the speed-up of the approximation, where we compared it to the `hypergeom` function from the Symbolic Math Toolbox of Matlab as well as the incomplete Gamma function, needed for the computation of the truncated power law with Gaussian modes. The results showed a dramatic decrease of computational costs for above approximation by several orders of magnitude (see Fig. 5b). By virtue of this approximation the evaluation of the truncated power-law spectrum with exponential modes had similar computational cost compared to Gaussian modes.

3.2. Quality of generated random fields

The quality of the numerically generated random fields was determined by virtue of the kurtosis as well as the variogram. In addition we investigated the quality not only with respect to the different methods but also depending on the Hurst coefficient as well as the respective spectrum. For the last task we considered three spectra in addition to the truncated power-law spectra with Gaussian and exponential modes: (i) the classic Gaussian spectrum, (ii) the classic exponential spectrum and (iii) the Kolmogorov spectrum.

3.2.1. Kurtosis

As a measure for the Gaussianity of the generated random fields we choose the kurtosis (see Section 2.3). Since the kurtosis of a Gaussian random variable is $G_4 = 3$ we could easily assess and compare the different methods by estimating this value from numerically generated realizations.

To that end we evaluated and subsequently averaged the Kurtosis of 2000 numerically generated random fields over a wide range of spatial scales. The procedure, for all four different methods described above, showed only minor differences between them (see Fig. 6). All four methods were therefore able to reproduce the desired Kurtosis with a comparable quality for the whole spatial domain.

It is interesting to note that deviations from the perfect value of $G_4 = 3$ occurred only for large values of x. This result was in contrast to the reproduction of the variogram. Although these deviations were low we saw minor differences, with the Fourier–Wavelet method giving the best and the Randomization and Hybrid method giving the worst results. The latter could be explained by the fact, that these two methods are constructed such that the second moment is reproduced most accurately. Higher moments, like the Kurtosis, might therefore not be represented as well.

3.2.2. Number of reproduced decades

In order to investigate if and how the quality of the generated fields increased with numerical costs (see Section 3.1) we used the estimator described in Section 2.3.3 with respect to the relevant numerical value described in Section 3.1, i.e. the number of the terms of the sum(s) used for the approximation of Equation (B.1) in every respective method.

In order to introduce the general behavior of the four different methods and to exemplify the procedure of estimating the number of reproduced decades we present arbitrary yet representative examples for each of the considered methods. To that end we show in analogy to Fig. 6 in Fig. 7 the estimated variogram for a single scenario. In contrast to Fig. 6 we could clearly see the differences between the four methods, with the Fourier method having the worst and the Hybrid and Fourier–Wavelet method having the best reproduction of the desired variogram. This qualitative analysis will be substantiated in the following by quantitative investigations.

The Fourier method. The Fourier method was the most simple of the four methods, which were considered in this study (see Section 2.2) and was included in the analysis in order to provide a benchmark.

The results showed a log-linear behavior for the quality with respect to numerical costs (see Fig. 8). A comparison of the scenarios with Gaussian and exponential modes showed a very similar behavior with respect to the slope of the curve.

The Randomization method. The results for the Randomization method also showed a log-linear behavior for the quality with respect to numerical costs (see Fig. 9). In comparison to the Fourier method we saw the same qualitative behavior but with a steeper slope, i.e. the quality increases faster with increasing effort (compare Fig. 8).

We also saw that the increase in quality was different for Gaussian and exponential modes, with the latter showing less reproduced decades. A clear difference to the Fourier method was the observation, that with higher N the estimated number of
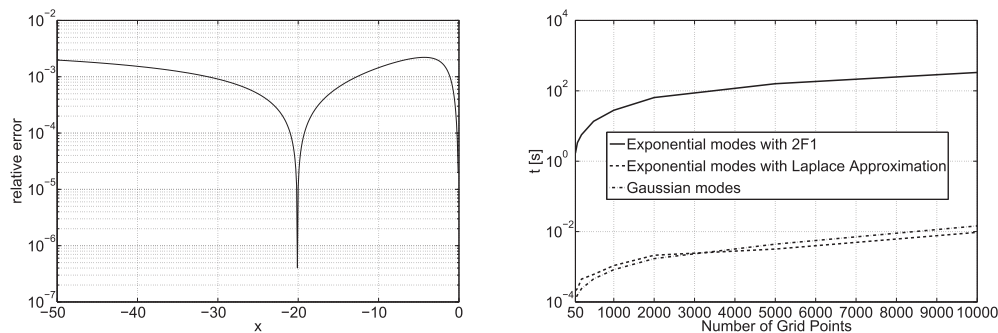
(a) Semi-log plot of the relative error $\varepsilon = | {}_2F_1 - {}_2\tilde{F}_1 | / {}_2F_1$. (b) Log-log plot of the computational costs of ${}_2F_1$ vs. ${}_2\tilde{F}_1$ and the incomplete Gamma function.
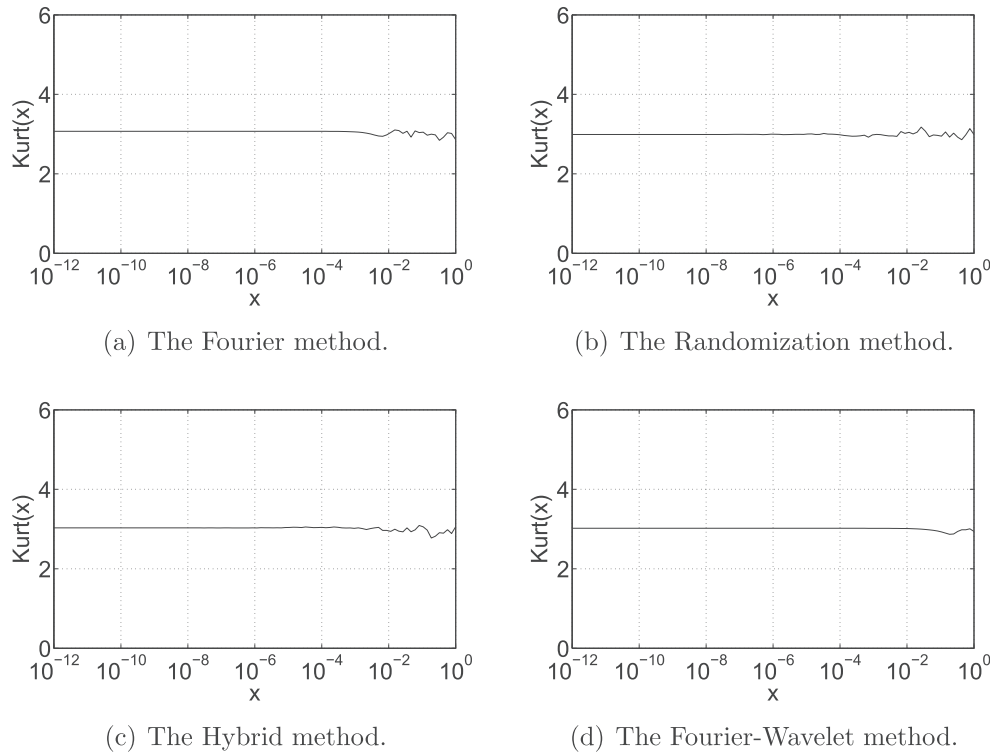
Fig. 5. Assessing the approximation of Gauss's hypergeometric function. The parameters were $H = 0.35$ and $n = 0.5$. (a) Semi-log plot of the relative error $\varepsilon = \left| {}_2F_1 - {}_2\tilde{F}_1 \right| / {}_2F_1$. (b) Log–log plot of the computational costs of ${}_2F_1$ vs. ${}_2\tilde{F}_1$ and the incomplete Gamma function.

Fig. 6. Estimated Kurtosis over several scales for an ensemble of 2000 realization. The parameters were $H = 0.35$ and $n = 0.5$ of a truncated power law with Gaussian modes. (a) The Fourier method. (b) The Randomization method. (c) The Hybrid method. (d) The Fourier−Wavelet method.

decades showed a high degree of variability. This could be explained by the random nature of the method itself. The problems of the Fourier method stem from the finite interval of the Fourier domain, which is sampled by a finite number of grid points N (see schematic in Fig. 2). In case of the Randomization method the interval, which is sampled for a given number of grid points, is not fixed but a random variable itself (see schematic in Fig. 3), which lead to random changes in the quality of the generated fields.
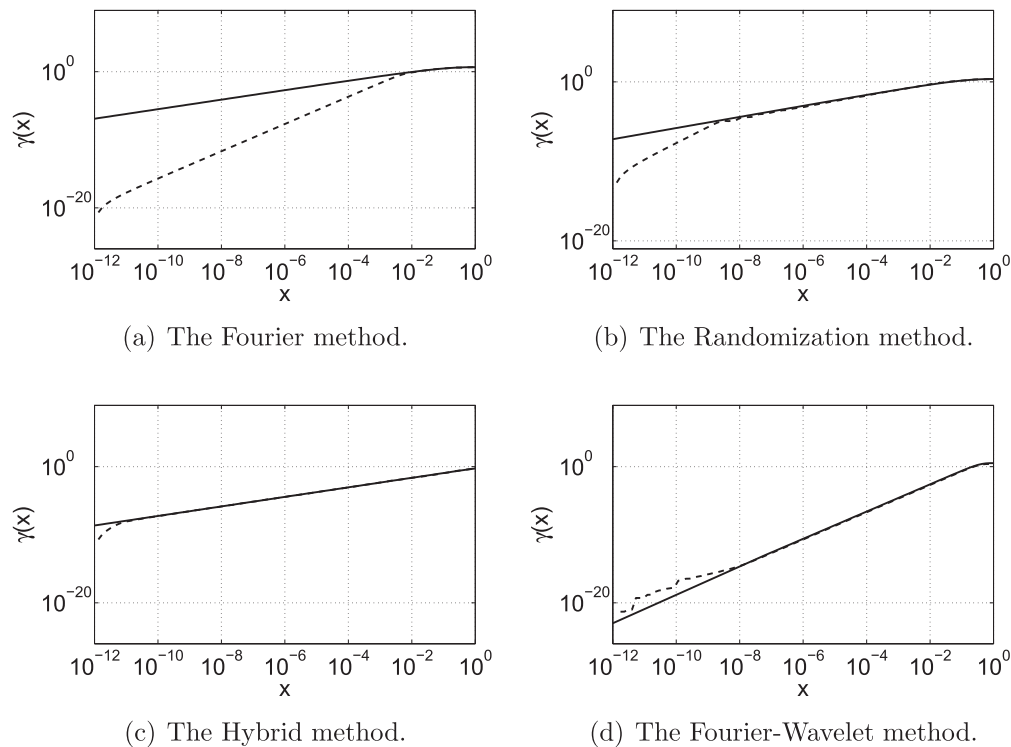

Fig. 7. Estimated Variogram (dashed line) vs. desired variogram (continuous line) over several scales for an ensemble of 2000 realization. The parameters were $H = 0.35$ and $n = 0.5$ of a truncated power law with Gaussian modes. (a) The Fourier method. (b) The Randomization method. (c) The Hybrid method. (d) The Fourier−Wavelet method.
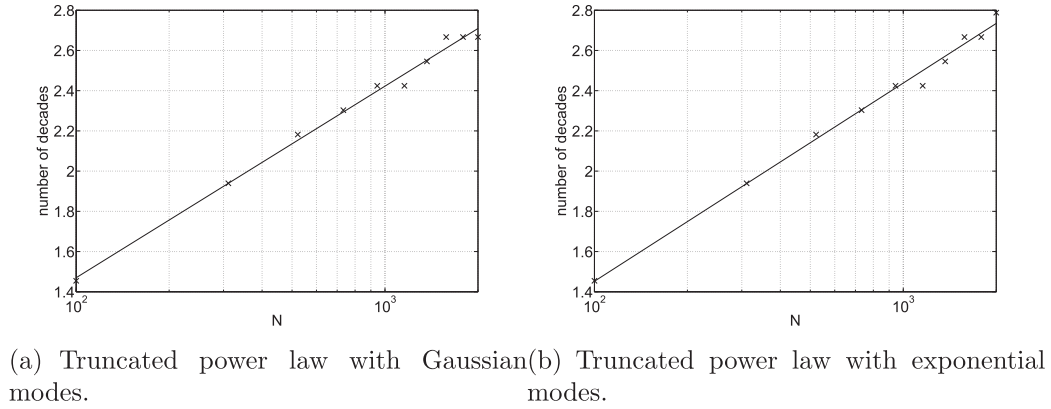
(a) Truncated power law with Gaussian modes.

(b) Truncated power law with exponential modes.

Fig. 8. Number of reproduced decades with respect to *N* (see Equation (17)) for the Fourier method for an ensemble of 2000 realization. The parameters were $H = 0.35$ and $n = 0.5$. (a) Truncated power law with Gaussian modes. (b) Truncated power law with exponential modes.
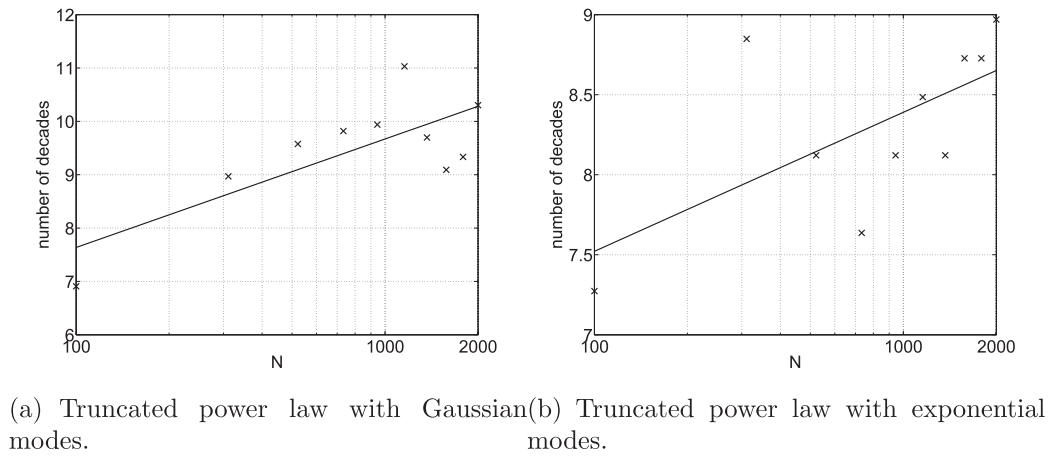

(a) Truncated power law with Gaussian modes.

(b) Truncated power law with exponential modes.

Fig. 9. Number of reproduced decades with respect to *N* (see Equation (22)) for the Randomization method for an ensemble of 2000 realization. The parameters were $H = 0.35$ and $n = 0.5$. (a) Truncated power law with Gaussian modes. (b) Truncated power law with exponential modes.

The Hybrid method. The results for the Hybrid method showed almost no dependency of the quality of the generated numerical fields with respect to the numerical costs (see Fig. 10).

This intriguing feature could be interpreted such that we get for very low numbers of *N* already the maximum reproduction of decades of the analytical variogram. In the simulations used for the results shown in Fig. 10 the number of partitions was $N_p = 40$. Neglecting the case of $N = 1$ we can therefore state that in this scenario the saturation was already reached by using the minimum number of 2 randomly chosen grid points $k_{i,j}$ per partition. It was furthermore observed, that, in conformance with the Fourier and the Randomization method, a different behavior for Gaussian as well as exponential modes prevailed.

The Fourier–Wavelet method. The results for the Fourier–Wavelet method showed a linear dependency of the quality of the generated numerical fields with respect to the numerical costs (see Fig. 11a) before eventually saturating in case of exponential modes (see Fig. 11b).

Similar to the other three investigated methods, we saw a different quality for Gaussian and exponential modes, with the former reaching up to 12 decades and the latter saturating at approximately 9 decades.

3.2.3. Quality depending on the spectrum

In the analysis above we saw for every considered method a consistent difference in the quality of the numerically generated random fields depending on the spectrum. This tendency was such that in case of Gaussian modes the quality was always better compared to exponential modes. As a result we investigated this tendency more thoroughly by considering three additional spectra: (i) a classic Gaussian spectrum as given by Equation (2b), (ii) a classic exponential spectrum as given by Equation (1b) as well as (iii) the Kolmogorov spectrum, which is given by Equation (4b).

Comparing the number of reproduced decades for these five spectra we could confirm the above mentioned tendency (see Table 1 for a single scenario), with a classic Gaussian spectrum giving the overall best and the truncated spectrum with exponential modes the overall worst results. The other spectra ranked in between with differences in quality depending on the numerical method.

The differences in quality could be explained with respect to a feature of said spectra called heavy tailing. A spectrum $S(k)$ is called heavy tailed if

$$\lim_{k \to \infty} e^{\lambda k}(1 - S(k)) = 0 \text{ for all } \lambda > 0 \tag{49}$$

holds. For the spectra in Table 1 this is true for the exponential, the Kolmogorov as well as the truncated (exponential) spectrum. The other two (Gaussian and truncated (Gaussian) spectrum) are short tailed. Heavy tailed spectra do not have a variance and the three spectra above do not even have an expectation value. Owing to that,
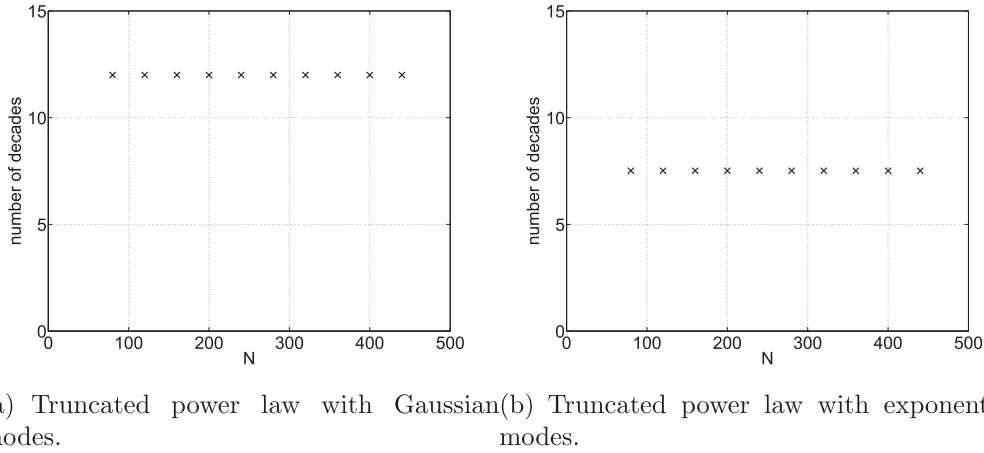
(a) Truncated power law with Gaussian modes. (b) Truncated power law with exponential modes.

Fig. 10. Number of reproduced decades with respect to $N \times N_p$ (see Equation (25)) for the Hybrid method for an ensemble of 2000 realization. The parameters were $N_p = 40$, $H = 0.35$ and $n = 0.5$. (a) Truncated power law with Gaussian modes. (b) Truncated power law with exponential modes.
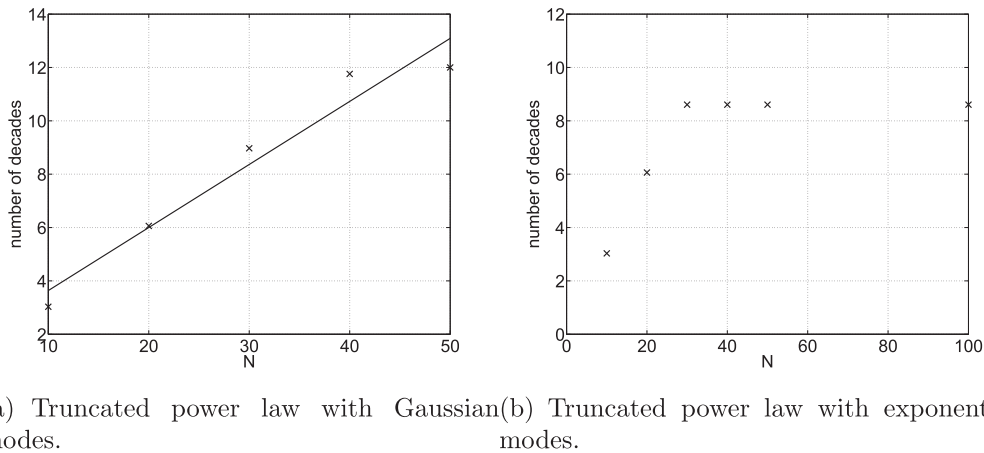

(a) Truncated power law with Gaussian modes. (b) Truncated power law with exponential modes.

Fig. 11. Number of reproduced decades with respect to N (see Equation (33)) for the Fourier–Wavelet method for an ensemble of 2000 realization. The parameters were $H = 0.35$ and $n = 0.5$. (a) Truncated power law with Gaussian modes. (b) Truncated power law with exponential modes.

it was much harder to represent the full characteristics of these spectra in the numerically generated random fields, which lead to the observed spectrum-dependent differences in quality.

This problem was strongest in case of a truncated power-law spectrum with exponential modes, which can be illustrated by the fact that for $H = 0.35$ and $\lambda_u = 2.0$ it takes an interval of $k \in [-2000 \ 2000]$ in order to sample 99% of the spectrum, whereas in case of the most benign Gaussian spectrum the same amount of mass is already concentrated within the interval of $k \in [-1.7 \ 1.7]$.

3.2.4. Quality depending on the Hurst coefficient

For the analysis above the Hurst coefficient was always kept constant at $H = 0.35$, in order to facilitate comparable results for the different methods and spectra. In addition we investigated how the quality of the numerically generated random fields depended on this coefficient.

To that end we ran several simulations with a varying Hurst coefficient H and applied subsequently the above method for estimating the number of reproduced decades. The results showed for both scenarios (Gaussian and exponential modes) that problems can occur for either very high or very low values of H (see Fig. 12). For intermediate values the performance was always comparable. Although the results shown in Fig. 12 cover only the Fourier–Wavelet method, our observations were similar for the other three

investigated methods (not shown here). According to data from the literature the value of H often lies in the midrange of the $0 \leq H \leq 0.5$ interval (Molz and Boman, 1995; Liu and Molz, 1996). The problems reported here may therefore be of lesser consequence for practical applications.

3.3. Structural analysis of random fields

In order to investigate the structural properties of the generated random fields we generated a data set of 100 realizations of two-dimensional random fields (see Appendix C). The generated random fields had a resolution of 1000×1000 grid points, i.e. 10^6 grid points in total. Due to computational considerations we used two-dimensional fields only. To the best of our knowledge no native extension of the Fourier–Wavelet method to higher dimensions exists. As a result we confined this analysis to the Fourier method, the Randomization method and the Hybrid method.

3.3.1. Comparison between generation methods

The first criterion, that was investigated was the mean chord length $\langle p_c \rangle$ of the clusters as a function of scale. Since the used domains were isotropic, sampling chords along horizontal and vertical test lines were sufficient for the characterization. In the fractal regime $\langle p_c \rangle$ normalized by scale was expected to be constant

Table 1
Number of reproduced decades for every considered spectrum. The number of grid points was $N = 1024$ and the Hurst-coefficient was $H = 0.35$ in case of a truncated power law.

Spectrum	Fourier	Random	Hybrid	FW
Gaussian	11	11	10	11
Truncated (Gaussian)	2	9	10	11.5
Exponential	3	6	11	11.5
Kolmogorov	3	6	11	11.5
Truncated (exponential)	2	8	8	8

since all morphological properties are scale-invariant, i.e. zooming in should always reveal new similar structures. Evidently, this condition could not be fulfilled for arbitrarily small scales and could therefore be used as a quality criterion for the different generation methods.

The reproduction of small-scale variation by the Fourier method was very poor in our tests, resulting in very regular clusters that comprise long chords (see Fig. 13). The generation of random fields by the Randomization method showed much better results, but the reproduction of small-scale variation was reproduced the best in case of the Hybrid method.

Like $\langle p_c \rangle$ the roughness of clusters, measured by virtue of the normalized boundary length L_B, should theoretically be scale-invariant in the fractal regime. A property, which was not fulfilled by any of the generation methods (see Fig. 14). Amongst these the Hybrid method performed best in preserving small-scale roughness of the clusters. The Fourier method failed completely in reproducing this small-scale variability, which is why the standard deviation between 100 realizations was also low. Note that the boundary length and mean chord length carry to some extent redundant information as they are related through stereological principles (Roberts and Torquato, 1999).

3.3.2. Comparison between Gaussian and exponential modes

Another aspect, where structural properties of a random field become relevant is the transition between the fractal and classic regime and possible differences between Gaussian and exponential modes. The numerical method used for the generation of random fields was the Hybrid method, since the above analysis showed the best overall performance of this method. The cut-off length was set to $\lambda_u = 1$ in all realizations. The maximum numerical length varied between $10^{-3} \leq L_{num} \leq 10^2$, i.e. 6 orders of scale were covered ranging from a purely fractal regime ($L_{num} \ll \lambda_u$), an intermediate regime ($L_{num} \approx \lambda_u$) to a classic regime ($L_{num} \gg \lambda_u$). The investigation of the last regime might seem superfluous since λ_u typically refers to the size of the investigated geological domain therefore

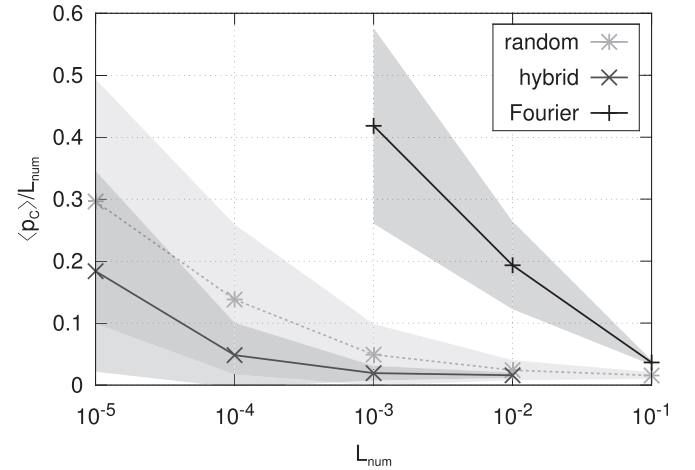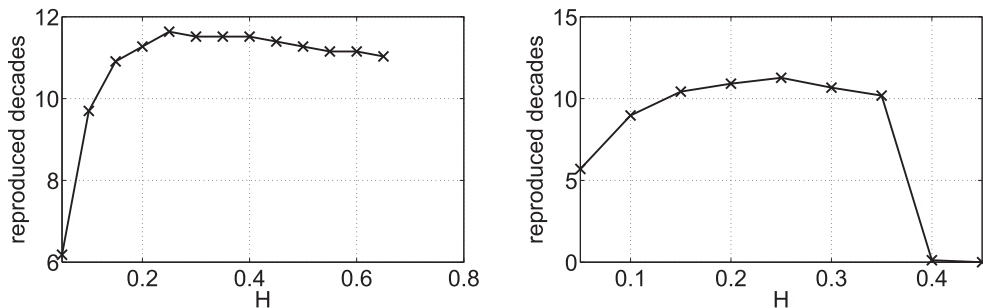**Fig. 13.** Mean chord length of the white phase normalized by scale $\langle p_c \rangle / L_{num}$ in a binary image thresholded at the mean gray value of each realization for random fields with different generation methods. Shaded areas represent standard deviation.

rendering the simulation of length scales bigger than λ_u unnecessary. We did however, include this regime into the analysis for two reasons. Firstly, in order to fully investigate the transition from a fractal to a classic regime. In addition, in case of a hybrid representation of a fractured medium, we do indeed have a situation, where $L_{num} \gg \lambda_u$. In such a scenario the large dominant fractures are modeled explicitly and are embedded within a continuous porous host medium, with the latter comprising the combined effects of the smaller background fractures. Here λ_u would no longer refer to the total size of the whole medium but to the much smaller size of the continuous compartments in between the discrete dominant fractures.

In order to investigate this transition quantitatively we used various morphological measures that assess spatial properties of the underlying structures by different means.

Roughness. Since boundary length and chord length carry similar information on the regularity of clusters we only present and discuss the results for the boundary length in the following.

The normalized boundary length remained rather constant within the fractal regime (L_{num} from 10^{-3} to 10^0) and increased abruptly as soon as scales with non-fractal character were used (L_{num} from 10^1 to 10^2) (see Fig. 15). Due to a higher inherent irregularity this threshold behavior was less distinct and L_B was higher for $L_{num} > 10^0$ in case of random fields with exponential modes than in random fields with Gaussian modes.

(a) Truncated power law with Gaussian modes.

(b) Truncated power law with exponential modes.

Fig. 12. Number of reproduced decades for the Fourier−Wavelet method with respect to the Hurst coefficient H (see Equation (33)) for a single numerical effort for an ensemble of 2000 realization. The parameters were $M = 40$ and $n = 0.5$. (a) Truncated power law with Gaussian modes. (b) Truncated power law with exponential modes.
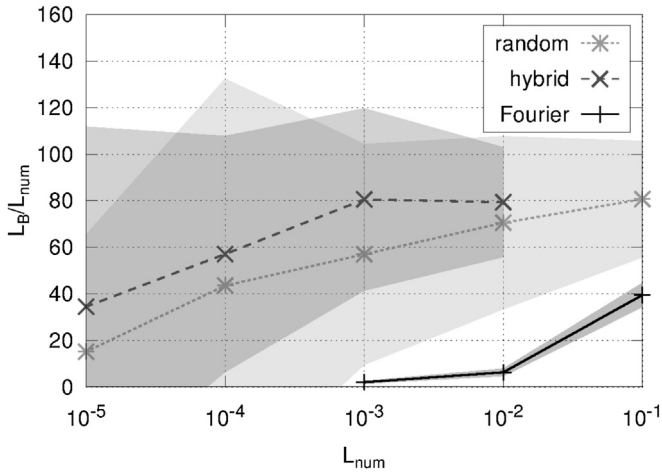
Fig. 14. Boundary length normalized by scale L_B/L_{num} in a binary image thresholded at the mean gray value of each realization for random fields with different generation methods. Shaded areas represent standard deviation.

Connectivity. The connectivity of a random field is a very important quantity with a high impact on the flow and transport behavior of a random field (Zinn and Harvey, 2003; Renard and Allard, 2011). Thus, different connectivities across scales or between random field models may provide an additional criterion to evaluate the suitability of a random field generator.

Results with respect to the two-point cluster function showed in general a high variability of the normalized connectivity length $\langle\tau\rangle$ between realizations, since the gray threshold for binarization coincided with the percolation threshold of the structure (see Fig. 16). In the fractal regime the mean connectivity length $\langle\tau\rangle$ was high due to the existence of large spanning clusters (see Fig. 4a). In the transition to the classic regime $\langle\tau\rangle$ decreased since these clusters were falling into pieces as the long-range correlation reduced. However, as the classical regime was fully reached $\langle\tau\rangle$ increased again, since a network of small, but connected clusters appeared. As a consequence, the variability among realizations also decreased since the long-range correlation had disappeared completely and the structure became translation-invariant.

Results with respect to the second connectivity measure, i.e. the normalized proportion of connected pairs of cluster pixels, showed
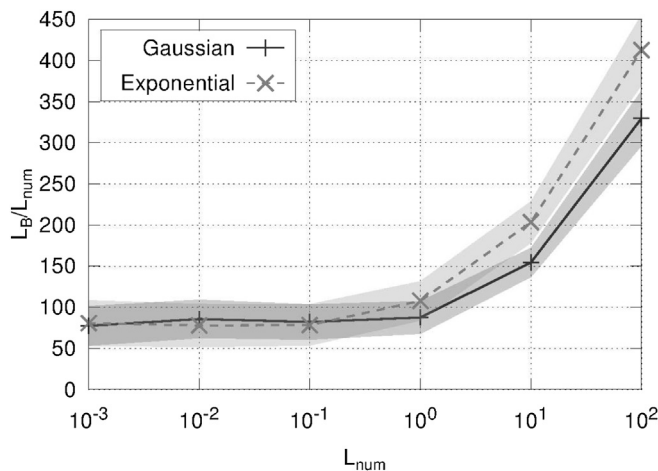
again that the transition from the fractal to the classic regime exhibited lower connectivity irrespective of mode (see Fig. 17). Just as $\langle\tau\rangle$, Γ too decreased stronger for exponential modes at the fractal end of the transition $L_{num} = 10^0$ compared to Gaussian modes. Hence, both measures have basically the same information content, at least for the investigated random fields. Again, the variability amongst realizations was quite high, since all level sets were at their percolation threshold.

3.3.3. Synthesis

These investigations showed that if L_{num} is smaller than λ_u a fractal regime prevails with Gaussian and exponential modes having essentially the same behavior. These findings are in complete agreement with what one would expect. It is in the transition to the classic regime for $L_{num} > \lambda_u$, where notable differences between both formulations of the truncated power law occurred. Due to the higher inherent irregularity, random fields with exponential modes formed clusters with lower thickness, higher surface roughness and partly lower connectivity.

4. Summary

In this paper we have tested and compared four different methods for the generation of random fields: (i) the Fourier method, (ii) the Randomization method, (iii) the Hybrid method and (iv) the Fourier−Wavelet method. Since no native extension of the Fourier−Wavelet method to two-dimensions exists we confined our analysis of higher dimensional properties to the first three methods. The variogram of the random fields was described by a truncated power law with both Gaussian and exponential modes. For one of the performed investigations we also considered additionally a classic Gaussian and exponential variogram as well as a pure power-law variogram (Kolmogorov variogram).

Several criteria were used in order to estimate the quality of random fields numerically generated by each of these four methods. The Kurtosis, as a measure for the Gaussianity of the field, and the variogram function were estimated from a large number of random field realizations (several thousands) and compared to the theoretically predicted values. Due to the large number of realizations only one-dimensional random fields were used. In addition we investigated the structural properties of the random fields by virtue of measures like roughness and connectivity. Since
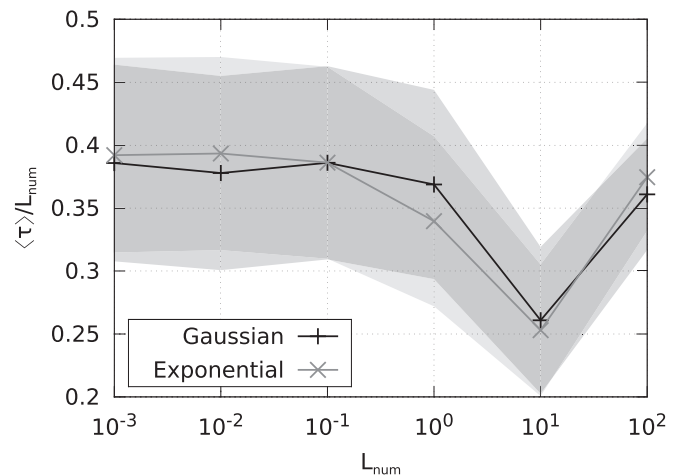

Fig. 15. Boundary length normalized by scale L_B/L_{num} in a binary image thresholded at the mean gray value of each realization for random fields within truncated power law and different modes. Shaded areas represent standard deviation.

Fig. 16. Mean connectivity length of objects normalized by scale $\langle\tau\rangle/L_{num}$ in a binary image thresholded at the mean gray value of each realization for random fields within Gaussian (dark gray) or exponential (light gray) modes. Shaded areas represent standard deviation.

Fig. 17. Second moment of cluster size distribution Γ of objects in level sets thresholded at the mean gray value of each realization for random fields within Gaussian (dark gray) or exponential (light gray) modes. Shaded areas represent standard deviation.
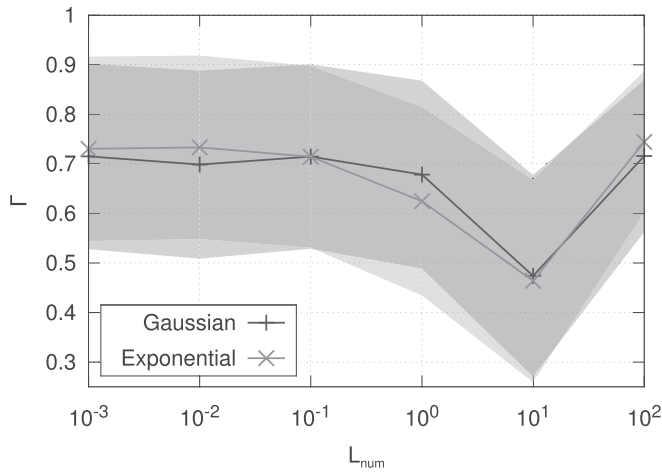
no analytical predictions could be used as comparison these measures were checked for plausibility. For the investigations of these measures two-dimensional random fields were used. The increased numerical costs were mitigated by the lower number of realizations that were used (100).

The results showed that all four methods were able to reproduce the Kurtosis over a wide range of scales with almost equal quality. This statement however, could not be extended to the reproduction of the variogram. Here strong differences between the different methods were found. The Fourier method performed worst, whereas the Hybrid and the Fourier–Wavelet method yielded the best results.

A similar tendency was found with respect to the development of computational cost. This development was measured as the number of reproduced scales of the estimated variogram versus the number of terms in the approximation of every method. Here the Fourier and the Randomization method showed a log-linear behavior, with the latter having a steeper slope. The Fourier–Wavelet method showed a linear behavior, i.e. the quality of the generated random fields was proportional to the applied numerical costs. The hybrid method showed basically no dependency between quality and numerical effort, i.e. this method is reaching the maximum performance already for a very small amount of terms in the approximation.

We furthermore investigated the quality of the numerically generated random fields with respect to the Hurst coefficient H in the truncated power law and with respect to the variogram itself. We found that the quality of the random fields suffered if very low (close to 0) of very high (close to 0.5 or 1 respectively) values for H was assumed. The quality for intermediate values did however, not differ significantly. Since extremely low and/or high values of H are of little practical relevance we could extend our above results, which were preformed for a single intermediate value of H, to a larger interval. In order to get more information on the possible influence of the variogram function itself on the quality of the numerical generation of random fields we considered three additional variograms: the classic Gaussian and exponential as well as the Kolmogorov variogram. The comparisons showed a clear tendency in the quality, with the classic Gaussian and the truncated power law with Gaussian modes giving consistently better results than the truncated power law with exponential modes. The classic exponential and the Kolmogorov variogram were of intermediate

quality. This tendency was consistently found for all of the four investigated methods. This behavior was explained by the heavy tailing of the associated spectra of each variogram. The more this heavy tailing was pronounced the harder it was to reproduce the variogram.

In the last part of this study we examined several structural properties of the numerically generated random fields on several scales. Since no native extension of the Fourier–Wavelet method to higher dimensions is known we used the Fourier method, the Randomization method and the Hybrid method, only. The results showed on overall similar trend compared to the reproduction of the variogram function, i.e. the Hybrid method showing the overall best, the Randomization method intermediate and the Fourier method the worst performance.

Acknowledgments

Appendix A. Random Fields

A mathematical representation of a random field is given by $\{u(\mathbf{x})|\mathbf{x}\in\Omega\}$. Here Ω is the support of u, i.e. the spatial region $\Omega\subseteq\mathbb{R}^d$ over which u is defined, with d being the spatial dimension ($d = 1,2,3$).

In order to fully describe a Gaussian random field it is sufficient to specify its expectation value as well as either its covariance or (semi)variogram function. The covariance of a random field $u(\mathbf{x})$ between two points is defined as

$$\mathrm{Cov}(u(\mathbf{x}_1), u(\mathbf{x}_2)) = E[(u(\mathbf{x}_1) - E[u(\mathbf{x}_1)])(u(\mathbf{x}_2) - E[u(\mathbf{x}_2)])], \tag{A.1}$$

with $\mathbf{x}_1$ and $\mathbf{x}_2$ being two arbitrary points in $u(\mathbf{x})$ and E being the expectation value. The variogram function (more precisely semivariogram or alternatively structure function according to some authors (Majda and Kramer, 1999; Cameron, 2003; Kramer et al., 2007)) is defined as

$$\gamma(u(\mathbf{x}_1), u(\mathbf{x}_2)) = \frac{1}{2}E\left[(u(\mathbf{x}_1) - u(\mathbf{x}_2))^2\right]. \tag{A.2}$$

Both these functions contain the same information in different forms, i.e. the degree of spatial dependency of a random field $u(\mathbf{x})$ between the points $\mathbf{x}_1$ and $\mathbf{x}_2$. It is therefore sufficient to use only one these in order to provide a full statistic characterization of the Gaussian random field $u(\mathbf{x})$. If not otherwise noted the variogram function is used in this study.

Appendix B. Spectral representation of Gaussian random fields

According to Mandelbrot and van Ness (1968) a Gaussian random field $u(x)$ can be represented by virtue of a stochastic Fourier integral

$$u(x) = \int_{-\infty}^{\infty} e^{-2\pi i k x}\sqrt{S(k)}\mathrm{d}W(k). \tag{B.1}$$

Here $\mathrm{d}W(k)$ is a complex-valued white noise random measure. This expression can be rewritten into

$$u(x) = \int_{-\infty}^{\infty} \cos(2\pi kx)\sqrt{S(k)}dW(k)$$

$$+ i\int_{-\infty}^{\infty} \sin(2\pi kx)\sqrt{S(k)}dW(k). \tag{B.2}$$

Since the spectrum $S(k)$ is the Fourier transform of the covariance function $Cov(h)$ of $u(x)$, which is an even function, $S(k)$ is even, too. As a result the following condition holds for the complex conjugate of the Wiener process $W(-k) = -\overline{W(k)}$. We therefore can rewrite Equation (B.2) as

$$u(x) = \sqrt{2\sigma^2}\left(\int_{0}^{\infty} \cos(2\pi kx)\sqrt{E(k)}dW^1(k)\right.$$

$$\left.+ \int_{0}^{\infty} \sin(2\pi kx)\sqrt{E(k)}dW^2(k)\right), \tag{B.3}$$

with $W^1(k)$ and $W^2(2)$ being two independent real valued Wiener processes and $E(k) = S(k)/\sigma^2$ being the normalized spectrum or spectral density function. Equation (B.3) is the final form, which is used for the derivation of spectral methods.

Appendix C. Extensions to higher dimensions

For the investigation how kurtosis and variogram function are reproduced, only one-dimensional random fields were used in this study. This decision was motivated by easier implementation and lower numerical costs and was seen adequate for these investigations. In order to investigate structural properties as well as for practical applications higher-dimensional fields are needed. Starting point of the derivation is again the spectral representation of a random field as given by Equation (B.1) but for higher dimensions

$$u(\mathbf{x}) = \int_{-\infty}^{\infty} e^{-2\pi i\mathbf{x}\cdot\mathbf{k}}\sqrt{S(\mathbf{k})}dW(\mathbf{k}). \tag{C.1}$$

In the following we will detail the discretization for two dimensions, only. The extensions to three dimensions will then be straight forward. Of the four methods considered here the first three (the Fourier method, the Randomization method and the Hybrid method) have native extensions to higher dimensions, whereas the Fourier−Wavelet method needs a special procedure.

The Fourier method. The extension of the Fourier method for higher dimensions is relatively simple. The discretization of the integral as well as the Wiener process yields the following equation for two dimensions

$$u(\mathbf{x})_{four} = \sqrt{2\sigma^2}\sum_{i=0}^{N_x}\sum_{i=0}^{N_y}\sqrt{E(\mathbf{k})}\left(Z_{i,j}^1 \cos(2\pi\mathbf{k}\cdot\mathbf{x})\right.$$

$$\left.+ Z_{i,j}^2 \sin(2\pi\mathbf{k}\cdot\mathbf{x})\right)\sqrt{\Delta k_x\Delta k_y} \tag{C.2}$$

Here Δk_x and Δk_y are the finite step size in the x and y direction, N_x and N_y are the respective number of grid points and $\mathbf{x} = (x_i,x_j)$ and $\mathbf{k} = (\mathbf{k_i},\mathbf{k_j})$ are the discretized support in the original as well as the Fourier domain. The extension to three dimensions is straight forward and not described here.

The Randomization method. The Randomization method in one dimension is given by Equation (22). The extension to higher dimension is given by

$$u(\mathbf{x})_{rand} = \sqrt{2\frac{\sigma^2}{N}\sum_{i=1}^{N}\left(Z_i^1 \cos(2\pi\mathbf{k}_i\mathbf{x}) + Z_i^2 \sin(2\pi\mathbf{k}_i\mathbf{x})\right)}. \tag{C.3}$$

Here the $\mathbf{k}_i = (k_{x,i},k_{y,i})$ are the two-dimensional or $\mathbf{k}_i = (k_{x,i},k_{y,i},k_{z,i})$ the three-dimensional random variable for every mode i. The PDF of this random variables is again the spectral density $E(\mathbf{k}) = S(\mathbf{k})/\sigma^2$, i.e. the normalized spectrum. The main challenge is therefore the generation of the two- or three-dimensional random variables. In two dimensions a possible solution is the transformation from Cartesian into polar coordinates, i.e. a representation like

$$\begin{aligned}k_x &= r\cos(2\pi\varphi),\\ k_y &= r\sin(2\pi\varphi).\end{aligned} \tag{C.4}$$

Here φ is a uniformly distributed random variable and r is a random variable distribute according to

$$E_r(r) = 2\pi rE(r), \tag{C.5}$$

where E is the two-dimensional spectral density and $r = \sqrt{k_x^2 + k_y^2}$ is the length of the random point in the two-dimensional Fourier domain. For three dimensions a similar procedure can be used, where spherical instead of polar coordinates are used

$$\begin{aligned}k_x &= r\sin(\theta)\cos(2\pi\varphi),\\ k_y &= r\sin(\theta)\sin(2\pi\varphi),\\ k_z &= r\cos(\theta).\end{aligned} \tag{C.6}$$

Here the azimuth φ is again a uniformly distribute random variable and the inclination θ is given as

$$\theta = \frac{\sin(2\pi\zeta)}{2}, \tag{C.7}$$

with ζ being a uniformly distribute random variable. The radius r is distributed according to

$$E_r(r) = 4\pi r^2 E(r), \tag{C.8}$$

This procedure is applicable since an anisotropic random field can be transformed into an isotropic random field.

The Hybrid method. The Hybrid method in one dimension is given by Equation (25). Due to the similarities between the Randomization and the Hybrid method we can extend the statements made above to this method, too. We are therefore not repeating the procedure here.

The Fourier−Wavelet method. The one-dimensional Fourier−Wavelet method given by Equation (33) can not easily be expanded to two or three dimensions. To that end we need to apply a special procedure. According to Elliott and Majda (1995) we can construct a two- or three-dimensional random field by superimposing one-dimensional random field with the desired properties

$$\mathbf{u}(\mathbf{x})_{fw} = \frac{1}{N}\sum_{i=1}^{N}[\mathbf{P}(\mathbf{r}_i)\mathbf{u}_i^{(1)}(\mathbf{r}_i\cdot\mathbf{x}) + (\mathbf{I} - \mathbf{P}(\mathbf{r}_i))\mathbf{u}_i^{(2)}(\mathbf{r}_i\cdot\mathbf{x})]. \tag{C.9}$$

Here $\mathbf{u}_i^{(1)}$ and $\mathbf{u}_i^{(2)}$ are mutually independent Gaussian d-dimensional random fields, each component of which is generated by the one-dimensional Fourier−Wavelet method as given by Equation (33) and $\mathbf{I}$ is the identity matrix. The projection matrix $\mathbf{P}$ is given by

$$P(\mathbf{k}) = \frac{k_ik_j}{|\mathbf{k}|}, \quad i,j = 1,\cdots,d. \tag{C.10}$$

By virtue of this matrix the vectors $\mathbf{r}_i$ are deterministically sampling the unit circle or unit sphere in two- or three dimensions, respectively.

It should be noted that this method can be applied for every other method in order to generate random fields with two or three dimensions. However, for the three other methods, investigated in this study, simpler and therefore preferable procedures exist. The Fourier—Wavelet method is the only method without a native extension so the decomposition given by Equation (C.9) is mandatory.

References

Abramowitz, M., Stegun, I., 1970. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, ninth ed. Dover Publications.

Allard, D., 1993. On the connectivity of two random set models: the truncated gaussian and the boolean. Quant. Geolo. Geostat. 5, 467—478.

Anderson, A., McBratney, A., Crawford, J., 1998. Applications of fractals to soil studies. In: Sparks, D.L. (Ed.), Advances in Agronomy, vol. 63, pp. 1—76.

Bear, J., 1972. Dynamics of Fluids in Porous Media. American Elsevier.

Bellin, A., Rubin, Y., Nov 1996. HYDRO_GEN: a spatially distributed random field generator for correlated properties. Stoch. Hydrol. Hydraul. 10 (4), 253—278.

Benson, D., Schumer, R., Meerschaert, M., Wheatcraft, S., Jan 2001. Fractional dispersion, Levy motion, and the MADE tracer tests. Transp. Porous Media 42 (1—2), 211—240.

Berkowitz, B., 2002. Characterizing flow and transport in fractured geological media: a review. Adv. Water Resour. 25 (8—12), 861—884.

Berkowitz, B., Cortis, A., Dentz, M., Scher, H., 2006. Modeling non-fickian transport in geological formations as a continuous time random walk. Rev. Geophys. 44.

Boggs, J., Young, S., Beard, L., Gelhar, L., Rehfeldt, K., Adams, E., Dec 1992. Field-study of dispersion in a heterogeneous aquifer .1. Overview and site description. Water Resour. Res. 28 (12), 3281—3291.

Bouchaud, J., Georges, A., Nov 1990. Anomalous diffusion in disordered Media — statistical mechanisms, models and physical applications. Phys. reports — Rev. Sect. Phys. Lett. 195 (4—5), 127—293.

Brouste, A., Istas, J., Lambert-Lacroix, S., Nov 2007. On fractional Gaussian random fields simulations. J. Stat. Softw. 23 (1), 1—23.

Butler, R.W., Wood, A.T.A., 2002. Laplace approximations for hypergeometric functions with matrix argument. Ann. Stat. 30 (4), 1155—1177.

Cameron, C., Dec 10 2003. Relative efficiency of Gaussian stochastic process sampling procedures. J. Comput. Phys. 192 (2), 546—569.

Caspi, A., Granek, R., Elbaum, M., Dec 25 2000. Enhanced diffusion in active intracellular transport. Phys. Rev. Lett. 85 (26, 1), 5655—5658.

Caspi, A., Granek, R., Elbaum, M., Jul 2002. Diffusion and directed motion in cellular transport. Phys. Rev. E 66 (1, 1).

Cirpka, O.A., Nowak, W., Feb 2003. Dispersion on kriged hydraulic conductivity fields. Water Resour. Res. 39 (39).

Coeurjolly, J.-F., 2000. Simulation and identification of the fractional brownian motion: a bibliographical and comparative study. J. Stat. Softw. 5 (7), 1—53.

Cohen, S., Istas, J., 2013. Fractional Fields and Applications. In: Mathmatiques et Applications, vol. 73. Springer.

Dagan, G., 1989. Flow and Transport in Porous Formations. Springer.

Dai, Z., Wolfsberg, A., Lu, Z., Reimus, P., Apr 14 2007. Upscaling matrix diffusion coefficients for heterogeneous fractured rocks. Geophys. Res. Lett. 34 (7).

Dentz, M., Kinzelbach, H., Attinger, S., Kinzelbach, W., Jul 2002. Temporal behavior of a solute cloud in a heterogeneous porous medium 3. numerical simulations. Water Resour. Res. 38 (7).

Deutsch, C.V., Journel, A.G., 1997. GSLIB: Geostatistical Software Library and User's Guide, second ed.. In: Applied Geostatistics Series Oxford University Press.

Di Federico, V., Neuman, S.P., May 1997. Scaling of random fields by means of truncated power variograms and associated spectra. Water Resour. Res. 33 (5), 1075—1085.

Dietrich, C., Newsam, G., AUG 1993. A fast and exact method for multidimensional gaussian stochastic simulations. Water Resour. Res. 29 (8), 2861—2869.

Dykaar, B., Kitanidis, P., Apr 1992. Determination of the effective hydraulic conductivity for heterogeneous Porous-Media using a numerical spectral Approach.1. Method. Water Resour. Res. 28 (4), 1155—1166.

Elliott, F.W., Majda, A.J., 1995. A new algorithm with plane waves and wavelets for random velocity fields with many spatial scales. J. Comput. Phys. 117 (1), 146—162.

Elliott, F., Horntrop, D., Majda, A., Apr 1997. A Fourier—wavelet Monte Carlo method for fractal random fields. J. Comput. Phys. 132 (2), 384—408.

Emery, X., Dec 2004. Testing the correctness of the sequential algorithm for simulating Gaussian random fields. Stoch. Environ. Res. Risk Assess. (SERRA) 18 (6), 401—413.

Franceschetti, G., Iodice, A., Migliaccio, M., Riccio, D., Sep 1999. Scattering from natural rough surfaces modeled by fractional Brownian motion two-dimensional processes. IEEE Trans. Antennas Propag. 47 (9), 1405—1415.

Gautschi, W., 2002. Gauss quadrature approximations to hypergeometric and confluent hypergeometric functions. J. Comput. Appl. Math. 139, 173—187.

Gil, A., Segura, J., Temme, N.M., 2006. The abc of hyper recursions. J. Comput. Appl. Math. 190, 270—286.

Gimenez, D., Perfect, E., Rawls, W., Pachepsky, Y., Dec 1 1997. Fractal models for predicting soil hydraulic properties: a review. Eng. Geol. 48 (3—4), 161—183.

Guadagnini, A., Riva, M., Neuman, S.P., 2012. Extended power-law scaling of heavy-tailed random air-permeability fields in fractured and sedimentary rocks. Hydrol. Earth Syst. Sci. 16 (9), 3249—3260.

Guadagnini, A., Neuman, S., Schaap, M., Riva, M., 2013. Anisotropic statistical scaling of soil and sediment texture in a stratified deep vadose zone near maricopa, arizona. Geoderma 214—215, 217—227.

He, Y., Hu, K.L., Huang, Y.F., Li, B.G., Chen, D.L., Jun 2010. Analysis of the anisotropic spatial variability and three-dimensional computer simulation of agricultural soil bulk density in an alluvial plain of north China. Math. Comput. Model. 51 (11—12), 1351—1356.

Hornung, U., 1996. Homogenization and Porous Media. Springer.

Hsu, K.-C., Chen, K.-C., Oct 2010. Multiscale flow and transport model in three-dimensional fractal porous media. Stoch. Environ. Res. Risk Assess. (SERRA) 24 (7), 1053—1065.

Huang, G., Huang, Q., Zhan, H., 2006. Evidence of one-dimensional scale-dependent fractional advection-dispersion. J. Contam. Hydrol. 85 (1—2), 53—71.

Hunt, J.C.R., Phillips, O.M., Williams, D., 1991. Turbulence and Stochastic Processes: Kolmogorov's Ideas 50 Years on. Royal Society.

Hurst, H., 1951. Long-term storage capacity of reservoirs. Trans. Am. Soc. Civ. Eng. 116, 770—799.

Hyun, Y., Neuman, S., Vesselinov, V., Illman, W., Tartakovsky, D., Di Federico, V., Jun 2002. Theoretical interpretation of a pronounced permeability scale effect in unsaturated fractured tuff. Water Resour. Res. 38 (6).

Ibrahim, A.K., Rakha, A.R., 2008. Contiguous relations and their computations for 2F1 hypergeometric series. Comput. Math. Appl. 56, 1918—1926.

Jiao, Y., Stillinger, F.H., Torquato, S., Oct 20 2009. A superior descriptor of random textures and its predictive capacity. Proc. Natl. Acad. Sci. U. S. A. 106 (42), 17634—17639.

Juanes, R., Tchelepi, H.A., Sep 2008. Special issue on multiscale methods for flow and transport in heterogeneous porous media. Comput. Geosci. 12 (3), 255—256.

Kramer, P.R., Kurbanmuradov, O., Sabelfeld, K., Sep 10 2007. Comparative analysis of multiscale Gaussian random field simulation algorithms. J. Comput. Phys. 226 (1), 897—924.

Kroese, D.P., Botev, Z.I., Aug 2013. Spatial Process Generation. ArXiv e-prints.

Liu, H.H., Molz, F.J., 1996. Discrimination of fractional brownian movement and fractional gaussian noise structures in permeability and related property distributions with range analyses. Water Resour. Res. 32 (8), 2601.

Liu, H.H., Zhang, Y.Q., Zhou, Q., Molz, F.J., Feb 20 2007. An interpretation of potential scale dependence of the effective matrix diffusion coefficient. J. Contam. Hydrol. 90 (1—2), 41—57.

Lovejoy, S., Schertzer, D., 2007. Scaling and multifractal fields in the solid earth and topography. Nonlinear Process. Geophys. 14 (4), 465—502.

Majda, A.J., Kramer, P.R., 1999. Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena. Phys. Reports 314 (4—5), 237—574.

Mandelbrot, B., van Ness, J., 1968. Fractional brownian motions, fractional noises and applications. SIAM Rev. 10 (4), 422.

Mantz, H., Jacobs, K., Mecke, K., Dec 2008. Utilizing Minkowski functionals for image analysis: a marching square algorithm. J. Stat. Mech. — Theory Exp. 12, 12015—12044.

Marty, R., Solna, K., Feb 2011. A general framework for waves in random Media with long-range correlations. Ann. Appl. Probab. 21 (1), 115—139.

Mecke, K., 2000. Statistical Physics and Spatial Statistics the Art of Analyzing and Modeling Spatial Structures and Pattern Formation. In: Lecture Notes in Physics, vol. 554. Springer, p. 111184. Ch. Additivity, Convexity and Beyond: Application of Minkowski Functionals in Statistical Physics.

Metzler, R., Klafter, J., Dec 2000. The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys. Reports — Rev. Sect. Phys. Lett. 339 (1), 1—77.

Molz, F.J., Boman, G.K., 1995. Further evidence of fractal structure in hydraulic conductivity distributions. Geophys. Res. Lett. 22 (18), 2545—2548.

Molz, F.J., Liu, H.H., Szulga, J., 1997. Fractional brownian motion and fractional gaussian noise in subsurface hydrology: a review, presentation of fundamental properties, and extensions. Water Resour. Res. 33 (10), 2273—2286.

Molz, F.J., Rajaram, H., Lu, S., Mar 6 2004. Stochastic fractal-based models of heterogeneity in subsurface hydrology: origins, applications, limitations, and future research questions. Rev. Geophys. 42 (1).

Neuman, S.P., Aug 1990. Universal scaling of hydraulic conductivities and dispersivities in geologic media. Water Resour. Res. 26 (8), 1749—1758.

Neuman, S.P., Mar 1 1994. Generalized scaling of permeabilities — validation and effect of support scale. Geophys. Res. Lett. 21 (5), 349—352.

Neuman, S.P., Jun 1995. On Advective transport in fractal permeability and Velocity-Fields. Water Resour. Res. 31 (6), 1455—1460.

Neuman, S.P., 2005. Trends, prospects and challenges in quantifying flow and transport through fractured rocks. Hydrogeol. J. 13 (1), 124—147.

Neuman, S., Di Federico, V., Sep 30 2003. Multifaceted nature of hydrogeologic scaling and its interpretation. Rev. Geophys. 41 (3).

Neuman, S.P., Tartakovsky, D.M., 2009. Perspective on theories of non-fickian transport in heterogeneous media. Adv. Water Resour. 32 (5), 670—680.

Neuman, S.P., Riva, M., Guadagnini, A., Feb 5 2008. On the geostatistical characterization of hierarchical media. Water Resour. Res. 44 (2).

Neuman, S.P., Guadagnini, A., Riva, M., Siena, M., 2013. Advances in Hydrogeology. Springer, New York. Ch. Recent Advances in Statistical and Scaling Analysis of Earth and Environmental Variables, pp. 1—15.

Nordbotten, J.M., 2009. Adaptive variational multiscale methods for multiphase flow in porous media. Multisc. Model. Simul. 7 (3), 1455—1473.

Ouchi, S., Matsushita, M., May 1992. Measurement of self-affinity on surfaces as a trial application of fractal geometry to landform analysis. Geomorphology 5 (1—2), 115—130.

O'Malley, D., Cushman, J.H., O'Rear, P., Mar 10 2012. On generating conductivity fields with known fractal dimension and nonstationary increments. Water Resour. Res. 48.

Pearson, J., 2009. Computation of Hypergeometric Functions. Master's thesis. Worcester College, University of Oxford.

Raghavan, R., Jul 1 2006. Some observations on the scale dependence of permeability by pumping tests. Water Resour. Res. 42 (7).

Renard, P., Allard, D., 2011. Connectivity metrics for subsurface flow and transport. Adv. Water Resour. 51, 168—196.

Ritzi, R.W., Dai, Z., Dominic, D.F., Rubin, Y.N., 2004. Spatial correlation of permeability in cross-stratified sediment with hierarchical architecture. Water Resour. Res. 40, 14—28.

Riva, M., Neuman, S.P., Guadagnini, A., Jan 2013. Sub-Gaussian model of processes with heavy-tailed distributions applied to air permeabilities of fractured tuff. Stoch. Environ. Res. Risk Assess. (SERRA) 27 (1), 195—207.

Roberts, A., Torquato, S., May 1999. Chord-distribution functions of three-dimensional random media: approximate first-passage times of Gaussian processes. Phys. Rev. E 59 (5, Part a), 4953—4963.

Ruan, F., McLaughlin, D., Apr 15 1998. An efficient multivariate random field generator using the fast Fourier transform. Adv. Water Resour. 21 (5), 385—399.

Rubin, Y., 2003. Applied Stochastic Hydrogeology. Oxford University Press, USA.

Sahimi, M., Tajer, S., Apr 2005. Self-affine fractal distributions of the bulk density, elastic moduli, and seismic wave velocities of rock. Phys. Rev. E 71 (4, 2).

SanchezVila, X., Carrera, J., Girardi, J., Aug 1996. Scale effects in transmissivity. J. Hydrol. 183 (1—2), 1—22.

Schlueter, S., Vogel, H.-J., Feb 2011. On the reconstruction of structural and functional properties in random heterogeneous media. Adv. Water Resour. 34 (2), 314—325.

Schulze-Makuch, D., Carlson, D., Cherkauer, D., Malik, P., Nov 1999. Scale dependency of hydraulic conductivity in heterogeneous media. Ground Water 37 (6), 904—919.

Schumer, R., Meerschaert, M.M., Baeumer, B., 2009. Fractional advection-dispersion equations for modeling transport at the earth surface. J. Geophys. Res. Earth Surf. 114 (F4).

She, Z., LeVeque, E., Jan 17 1994. Universal scaling laws in fully-developped turbulence. Phys. Rev. Lett. 72 (3), 336—339.

Shlesinger, M., West, B., Klafter, J., Mar 16 1987. Lévy dynamics of enhanced diffusion — application to turbulence. Phys. Rev. Lett. 58 (11), 1100—1103.

Shlesinger, M., Zaslavsky, G., Klafter, J., May 6 1993. Strange kinetics. Nature 363 (6424), 31—37.

Siena, M., Guadagnini, A., Riva, M., Neuman, S.P., 2012. Extended power-law scaling of air permeabilities measured on a block of tuff. Hydrol. Earth Syst. Sci. 16 (1), 29—42.

Torquato, S., 2002. Random Heterogeneous Materials: Microstructure and Macroscopic Properties, Interdisciplinary Applied Mathematics, vol. 16. Springer, New York.

Torquato, S., Beasley, J., Chiew, Y., 1988. Two-point cluster function for continuum percolation. J. Chem. Phys. 88 (10), 65406547.

Trangenstein, J.A., 2002. Multi-scale iterative techniques and adaptive mesh refinement for flow in porous media. Adv. Water Resour. 25 (812), 1175—1213.

Vogel, H.J., Weller, U., Schlüter, S., Oct 2010. Quantification of soil structure based on Minkowski functions. Comput. Geosci. 36 (10), 1236—1245.

Western, A.W., Blöschl, G., Grayson, R.B., Jan 2001. Toward capturing hydrologically significant connectivity in spatial patterns. Water Resour. Res. 37 (1), 8397.

Whitaker, S., 1999. The Method of Volume Averaging. In: Theory and Applications of Transport in Porous Media. Kluwer, Dordrecht.

Zhou, Q., Liu, H.-H., Molz, F.J., Zhang, Y., Bodvarsson, G.S., Aug 15 2007. Field-scale effective matrix diffusion coefficient for fractured rock: results from literature survey. J. Contam. Hydrol. 93 (1—4), 161—187.

Zinn, B., Harvey, C., Mar 12 2003. When good statistical models of aquifer heterogeneity go bad: a comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. Water Resour. Res. 39 (3).