Simulation and analysis of scalable non-Gaussian statistically anisotropic random functions

Monica Riva [a,b,*], Marco Panzeri [a], Alberto Guadagnini [a,b], Shlomo P. Neuman [b]

[a] Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Piazza L. Da Vinci 32, 20133 Milano, Italy
[b] Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, USA

Available online 2 July 2015

1. Introduction

Many earth and environmental (as well as other) variables, Y, and their spatial or temporal increments, ΔY, exhibit non-Gaussian statistical scaling. Our own experience shows this to be specifically true for quantities controlling subsurface flow and transport such as soil textural composition, log permeability, porosity and unsaturated flow parameters. Previously we (Siena et al., 2012, 2014; Riva et al., 2013a, b; Guadagnini et al., 2012, 2013, 2014, 2015) were able to capture some key aspects of such scaling by treating Y or ΔY as standard sub-Gaussian random func-tions $Y(x) = UG(x)$ (e.g., Samorodnitsky and Taqqu, 1994) in which x is a (spatial or temporal) coordinate, $G(x)$ is a scalable zero-mean stationary Gaussian random function and U is a non-negative random variable independent of $G(x)$. A major attraction of this model is that all multivariate moments of Y and ΔY (that exist) are fully defined by the distribution of U and the first two (one- and two-point) moments of G. The model, however, did not allow us to reconcile two seemingly contradictory observations, namely that whereas sample frequency distributions of Y (or its logarithm) exhibit relatively mild non-Gaussian peaks and tails, those of ΔY display peaks that grow sharper and tails that become heavier with decreasing separation distance or lag. Variables reported to exhibit one or both of these behaviors include (but are not limited to) permeability (Painter, 1996; Riva et al., 2013a, b) and hydraulic conductivity (Liu and Molz, 1997; Meerschaert et al., 2004; Guadagnini et al., 2013), porosity (Painter, 1996; Guadagnini et al., 2014, 2015), electrical resistivity (Painter, 2001; Yang et al., 2009), soil and sediment texture (Guadagnini et al., 2014), sediment transport rate (Ganti et al., 2009), rainfall (Kumar and Foufoula-Georgiou, 1993), measured and simulated turbulent fluid velocity (Castaing, 1990; Boffetta et al., 2008), and magnetic fluctuation (von Papen et al., 2014). Our analysis is thus potentially relevant to subsurface as well as surface hydrology, geophysics, soil physics and a range of other disciplines.

The standard sub-Gaussian model fails to capture these behaviors consistently because (*a*) multiplying any one realization of

* Corresponding author at: Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Piazza L. Da Vinci 32, 20133 Milano, Italy. Tel.: +39 02 2399 6214.

E-mail address: monica.riva@polimi.it (M. Riva).

$G(x)$ by a random number U renders all realizations of Y and ΔY Gaussian, and (b) setting $\Delta Y = U\Delta G$ fails to ensure that both ensemble and sample distributions of ΔY scale with lag in the aforementioned manner. To overcome this Neuman et al. (2013) and Nan et al. (2015) considered a synthetic sample of $Y(x) = UG(x)$, or of its increments, to consist of many Gaussian real-izations having random standard deviations U, also known as a Gaussian scale mixture. Such synthetic samples are, however, difficult to reconcile with real data.

Recently Riva et al. (2015) was able to overcome the above difficulty by developing a new generalized sub-Gaussian model that reconciles the statistical behaviors of a random function and its increments. These authors (a) derived analytical expressions for bivariate and marginal distributions of Y and marginal distributions of ΔY, (b) explored their results on synthetically generated one-dimensional realizations of Y, and (c) proposed new approaches to infer key parameters characterizing the probability density functions (pdfs) of Y and ΔY based on samples of one or (more importantly and reliably) both of these variables.

Here we extend the generalized sub-Gaussian model of Riva et al. (2015) to multiple dimensions, present an algorithm to generate corresponding random realizations of statistically isotropic or anisotropic sub-Gaussian functions and illustrate it in two dimensions. We demonstrate the accuracy of our algorithm by comparing ensemble statistics of Y and ΔY (such as, mean, variance, variogram and probability density function) with those of Monte Carlo generated realizations. We end by exploring the feasibility of estimating all relevant parameters of our model by analyzing jointly spatial moments of Y and ΔY obtained from a single realization of Y.

2. Generalized sub-Gaussian model

Let the variable of interest be a stationary random function $Y(\boldsymbol{x}) = \langle Y\rangle + Y'(\boldsymbol{x})$ defined on a continuum of points, $\boldsymbol{x}$, in multidimensional Euclidean space where $\langle Y\rangle$ is a constant ensemble mean (expectation) and $Y'(\boldsymbol{x})$ a zero-mean random fluctuation about $\langle Y\rangle$. We represent $Y'(\boldsymbol{x})$ in the generalized sub-Gaussian form introduced by Riva et al. (2015) as

$$Y'(\boldsymbol{x}) = U(\boldsymbol{x})G(\boldsymbol{x}) \tag{1}$$

where $G(\boldsymbol{x})$ is a zero-mean stationary, isotropic or anisotropic random Gaussian function and $U(\boldsymbol{x})$ a random function of mutually independent, identically distributed (iid) non-negative random values independent of $G(\boldsymbol{x})$. Admissible probability distributions of $U(\boldsymbol{x})$ include (but are not limited to) α-stable, log-normal, exponential, Weibull and gamma. Riva et al. (2015) derived the statistical properties of $Y'(\boldsymbol{x})$ considering $U(\boldsymbol{x})$ to be log-normally distributed according to $U \equiv \ln N(0,(2-\alpha)^2)$ where $\alpha < 2$. In this case, the marginal probability density function (pdf) of $Y'(\boldsymbol{x})$ coincides with the Normal Log-Normal distribution, NLN, (e.g., Guadagnini et al., 2015 and references therein)

$$f_{Y'}(y') = \frac{1}{2\pi(2-\alpha)}\int_0^\infty \frac{1}{u^2} e^{-\frac{1}{2}\left(\frac{1}{(2-\alpha)^2}\ln^2\frac{u}{\sigma_G}+\frac{y'^2}{u^2}\right)} du \tag{2}$$

where σ_G is the standard deviation of G. As (2) is symmetric, odd-order moments of Y' vanish while the variance and kurtosis of Y' become, respectively,

$$\sigma_Y^2 = \langle Y'^2\rangle = e^{2(2-\alpha)^2}\sigma_G^2 \tag{3}$$

$$\langle Y'^4\rangle = 3e^{8(2-\alpha)^2}\sigma_G^4 \tag{4}$$

The pdf of increments, $\Delta Y(s = |\boldsymbol{x}_A - \boldsymbol{x}_B|) = Y(\boldsymbol{x}_A) - Y(\boldsymbol{x}_B) = U(\boldsymbol{x}_A)G(\boldsymbol{x}_A) - U(\boldsymbol{x}_B)G(\boldsymbol{x}_B)$ in which s is separation distance (lag) between locations $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ in multidimensional space, is

$$f_{\Delta Y}(\Delta y) = \frac{1}{2\pi^2\sqrt{2\pi}(2-\alpha)^2} \times \int_0^\infty\int_0^\infty \frac{e^{-\frac{1}{2}\left[\frac{1}{(2-\alpha)^2}\left(\ln^2\frac{u_1}{\sigma_G}+\ln^2\frac{u_2}{\sigma_G}\right)+\frac{(\Delta y)^2}{u_1^2+u_2^2-2u_1u_2\rho_G}\right]}}{u_2u_1\sqrt{u_1^2+u_2^2-2u_1u_2\rho_G}} du_2du_1 \tag{5}$$

where ρ_G is the coefficient of correlation between $G(\boldsymbol{x}_A)$ and $G(\boldsymbol{x}_B)$. As (5) differs from the NLN distribution (2), the distributions of increments ΔY corresponding to various lags differ from the distribution of their parent variable, Y. From (5) one derives lead statistical moments of ΔY. Once again moments of odd order vanish due to symmetry while the variance and kurtosis of ΔY become, respectively,

$$\langle \Delta Y^2\rangle = 2\sigma_G^2 e^{(2-\alpha)^2}[e^{(2-\alpha)^2} - \rho_G] \tag{6}$$

$$\langle \Delta Y^4\rangle = 6\sigma_G^4 e^{4(2-\alpha)^2}[e^{4(2-\alpha)^2} + 1 - 4e^{(2-\alpha)^2}\rho_G + 2\rho_G^2] \tag{7}$$

The distributions of Y' and ΔY tend toward the Gaussian as $\alpha \to 2$. Otherwise $f_{Y'}$ is leptokurtic and $f_{\Delta Y}$ scales with lag because so does its parameter ρ_G, the correlation function of G. Riva et al. (2015) have shown that $f_{\Delta Y}$ has sharper peaks and heavier tails at small lags than does $f_{Y'}$ when $\alpha > 2 - \sqrt{\ln 3} \approx 0.95$, the opposite being true at large lags. In particular, when $\alpha \geqslant 1.8$ the asymptotic value of excess kurtosis of increments $(\langle \Delta Y^4\rangle/\langle \Delta Y^2\rangle^2 - 3)$ at large lags becomes much smaller than 1, rendering $f_{\Delta Y}$ virtually Gaussian.

The variogram and covariance of Y are given, respectively, by

$$\gamma_Y = \frac{\langle \Delta Y^2\rangle}{2} = \sigma_G^2 e^{(2-\alpha)^2}(e^{(2-\alpha)^2} - \rho_G) = \sigma_G^2 e^{(2-\alpha)^2}(e^{(2-\alpha)^2} - 1) + \gamma_G e^{(2-\alpha)^2} \tag{8}$$

$$C_Y = \sigma_Y^2 - \gamma_Y = e^{(2-\alpha)^2} C_G \tag{9}$$

where γ_G and $C_G = \sigma_G^2\rho_G$ are the variogram and covariance of G, respectively. The variogram γ_Y is seen to include a nugget effect (constant value independent of lag) that vanishes in the Gaussian limit as $\alpha \to 2$. Based on (9) one finds the (directional) integral scale of Y to be $I_Y = e^{-(2-\alpha)^2} I_G$ where I_G is the corresponding integral scale of G. The latter shows that a lognormal subordinator dampens, but does not destroy, the effect that G has on the covariance structure of Y; the more does α deviate from its Gaussian limit of 2 the shorter is the (integral) correlation scale of Y.

3. Generation and analysis of synthetic functions

To illustrate the manner in which one generates synthetic realizations of random functions corresponding to our generalized sub-Gaussian model we consider a two-dimensional square grid of $N_e = 300 \times 300 = 90{,}000$ squares measuring $10^{-2} \times 10^{-2}$ in arbitrary consistent length units. We generate $N_{MC} = 50{,}000$ synthetic realizations of $Y'(\boldsymbol{x})$, $\boldsymbol{x} \equiv \{x_1, x_2\}$, on this grid according to (1) in two steps. In the first step we use a version of SGSIM (Deutsch and Journel, 1998) modified by us to generate a two-dimensional zero-mean stationary and anisotropic Gaussian random function, $G(\boldsymbol{x})$, constituting truncated fractional Brownian motion (tfBm; Di Federico et al., 1999). The latter is an infinite hierarchy of mutually uncorrelated, stationary and anisotropic random functions or modes each of which is characterized by a (in our example expo-nential) variogram having anisotropy ratio $a = \lambda_{x_2}/\lambda_{x_1}$ where λ_{x_1} and λ_{x_2} are principal integral scales parallel, respectively, to the x_1 and x_2 coordinates. The hierarchy is characterized by a truncated power variogram (TPV)

$$\gamma_G^2(\tilde{s}) = \gamma^2(\tilde{s};\lambda_u) - \gamma^2(\tilde{s};\lambda_l) \tag{10}$$

in which

$$\gamma^2(\tilde{s};\lambda_m) = \frac{A\lambda_m^{2H}}{2H}\left[1 - e^{-\frac{\tilde{s}}{\lambda_m}} + \left(\frac{\tilde{s}}{\lambda_m}\right)^{2H}\Gamma\left(1-2H,\frac{\tilde{s}}{\lambda_m}\right)\right] \quad m = l,u \tag{11}$$

A is a coefficient; H is a Hurst scaling exponent; λ_l and λ_u are lower and upper cutoff scales proportional, respectively, to the resolution and sampling domain scales of given data; and $\tilde{s} = \sqrt{s_{x_1}^2 + s_{x_2}^2/a^2}$ is lag in an equivalent isotropic domain of $G(\boldsymbol{x})$ where s_{x_1} and s_{x_2} are principal lags in the original anisotropic domain. In the limits as $\lambda_l \to 0$ (corresponding to point data) and $\lambda_u \to \infty$ (infinite sampling domain) $G(\boldsymbol{x})$ becomes (nonstationary) fBm. The covariance and correlation coefficient of $G(\boldsymbol{x})$ are given respectively by

$$C_G(\tilde{s}) = C(\tilde{s};\lambda_u) - C(\tilde{s};\lambda_l); \quad \rho_G(\tilde{s}) = C_G(\tilde{s})/\sigma_G^2 \tag{12}$$

where

$$\sigma_G^2 = \frac{A}{2H}(\lambda_u^{2H} - \lambda_l^{2H}) \tag{13}$$

$$C^2(\tilde{s};\lambda_m) = \frac{A\lambda_m^{2H}}{2H}\left[e^{-\frac{\tilde{s}}{\lambda_m}} - \left(\frac{\tilde{s}}{\lambda_m}\right)^{2H}\Gamma\left(1-2H,\frac{\tilde{s}}{\lambda_m}\right)\right] \tag{14}$$

Based on (12) one finds the principal integral scales of $G(x)$ to be

$$I_{Gx_1} = \frac{2H}{1+2H}\frac{\lambda_u^{1+2H} - \lambda_l^{1+2H}}{\lambda_u^{2H} - \lambda_l^{2H}}; \quad I_{Gx_2} = I_{Gx_1}a \tag{15}$$

Neuman (2003) provided a hydrogeologic rationale for (10)–(15) by demonstrating mathematically that they follow directly from a representation of the subsurface as a juxtaposition of discrete materials the attributes of which constitute random fields. As shown by Neuman, (10)–(15) are obtained in the limit as these materials overlap to form a multiple continuum (analogous to the widely employed but much simpler dual porosity and dual permeability continua). Our realizations below are performed with $A = 1$, $H = 0.33$, $\lambda_l = 10^{-2}$ and $\lambda_u = 1$ which render $\sigma_G = 1.20$. Two sets of realizations are generated, one isotropic with $a = 1$ and $I_{Gx_1} = I_{Gx_2} = 0.42$, the other anisotropic with $a = 0.5$, $I_{Gx_1} = 0.42$ and $I_{Gx_2} = 0.21$.

In the second step we generate a realization of $Y'(\boldsymbol{x})$ by multiplying each discrete value of $G(\boldsymbol{x})$ at grid nodes by a random log-normal draw of $U(\boldsymbol{x})$. We do so twice, once by setting $\alpha = 1.8$ to yield $\sigma_Y = 1.25$ and $I_{Yx_j}/I_{Gx_j} = 0.96$ ($j = 1, 2$), I_{Yx_j} being principal integral scales of Y and once by setting $\alpha = 1.5$ to result in $\sigma_Y = 1.54$ and $I_{Yx_j}/I_{Gx_j} = 0.78$. We thus obtain four test cases, labeled TC1–TC4, the input parameters of which are summarized in Table 1.

Table 1
Input variables, ensemble variance $\langle Y'^2\rangle$, minima/maxima of sample mean $\langle Y'\rangle_{N_{MC}j}$ and variance $\langle Y'^2\rangle_{N_{MC}j}$ at grid nodes j, minima/maxima as well as mean and variance of spatial mean M_{1i}^Y and variance M_{2i}^Y across realizations i, in each test case.

Input/Ensemble values			Sample moments		Single realization spatial moments	
Test case	a; α	$\langle Y'^2\rangle$ Eq. (3)	$\langle Y'\rangle_{N_{MC}j}\times 10^{-2}$ min; max	$\langle Y'^2\rangle_{N_{MC}j}$ min; max	M_{1i}^Y min; max; mean; var	M_{2i}^Y min, max, mean, var
TC1	1.0; 1.8	1.55	−2.17; 1.79	1.49; 1.59	-1.63; 1.68; -0.17×10^{-2}; 0.13	0.83; 3.31; 1.40; 0.06
TC2	1.0; 1.5	2.36	−2.55; 2.34	2.22; 2.47	-1.81; 1.87; -0.19×10^{-2}; 0.17	1.27; 5.07; 2.16; 0.14
TC3	0.5; 1.8	1.55	−2.23; 2.20	1.49; 1.58	-1.17; 1.15; 5.02×10^{-4}; 0.08	0.93; 3.21; 1.46; 0.04
TC4	0.5; 1.5	2.36	−2.81; 2.80	2.23; 2.50	-1.30; 1.29; 5.41×10^{-4}; 0.10	1.43; 5.03; 2.24; 0.09

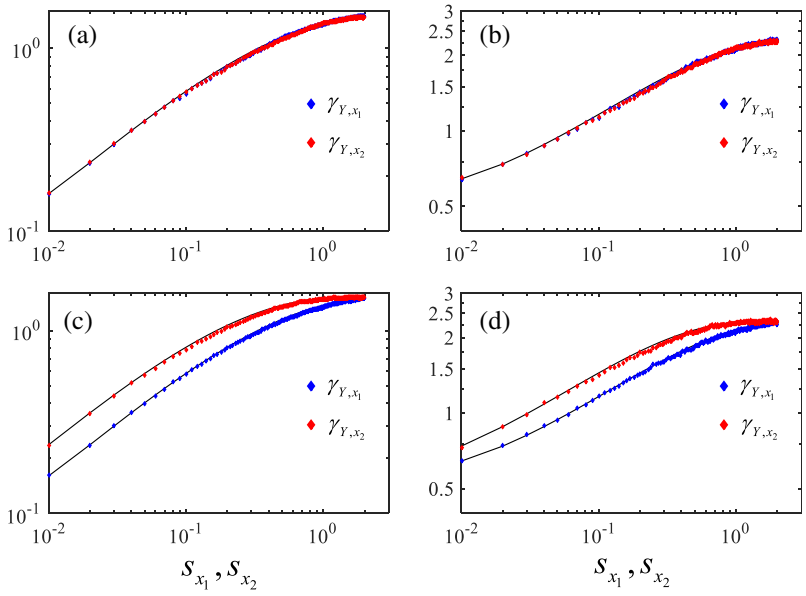

Fig. 1. Sample variograms γ_{Yx_1} and γ_{Yx_2} of Y' (symbols) along principal directions x_1 and x_2, respectively, at reference point $\boldsymbol{x} = (1,1)$ compared with theoretical variogram (solid curve) in (8) for (a) TC1, (b) TC2, (c) TC3, (d) TC4.

3.1. Sample versus ensemble statistics

We test the accuracy of our generation algorithm by comparing sample statistics of all $N_{MC} = 50{,}000$ realizations in each test case with their ensemble (theoretical) counterparts. The sample mean, $\langle Y' \rangle_{N_{MC}j}$, and variance, $\langle Y'^2 \rangle_{N_{MC}j}$, of Y' vary little from one grid node to another in each test case. As shown in Table 1, minimum and maximum values of $\langle Y' \rangle_{N_{MC}j}$ are on the order of 10^{-2}, close to the ensemble value of $\langle Y' \rangle = 0$, and those of $\langle Y'^2 \rangle_{N_{MC}j}$ differ from $\langle Y'^2 \rangle$ in (3) by less than 6% in all test cases.

Fig. 1 compares sample variograms γ_{Y,x_1} and γ_{Y,x_2} of Y' (symbols) in principal directions x_1 and x_2, respectively, at reference point $\boldsymbol{x} = (1,1)$ compared with theoretical variogram (solid curve) in (8) for the four test cases. Fig. 2 compares sample frequency distributions (symbols) of Y' at reference point $\boldsymbol{x} = (1.5, 1.5)$, and of ΔY at three lags parallel to x_2, starting from the same reference point, with corresponding theoretical pdfs (2) and (5), respectively, for all test cases. Fig. 2 clearly illustrates the way in which frequency distributions and pdfs of ΔY scale with lag. In both figures, the correspondence between sample and ensemble results is excellent. Results of similarly high quality were obtained for other choices of reference point and lag (not shown).

3.2. Single realization spatial statistics and parameter estimation

Next we compare lead spatial statistics of Y' and ΔY values, sampled across grid nodes within each individual realization

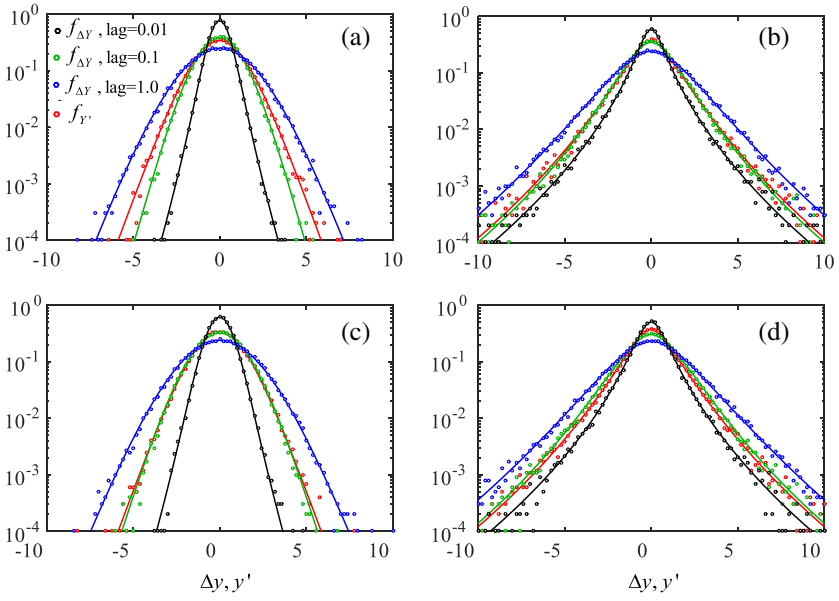

Fig. 2. Comparison of sample frequency distributions (symbols) of Y' at reference point $\boldsymbol{x} = (1.5, 1.5)$, and of ΔY at three lags parallel to x_2 starting from same reference point, with corresponding theoretical pdfs for (a) TC1, (b) TC2, (c) TC3, (d) TC4.

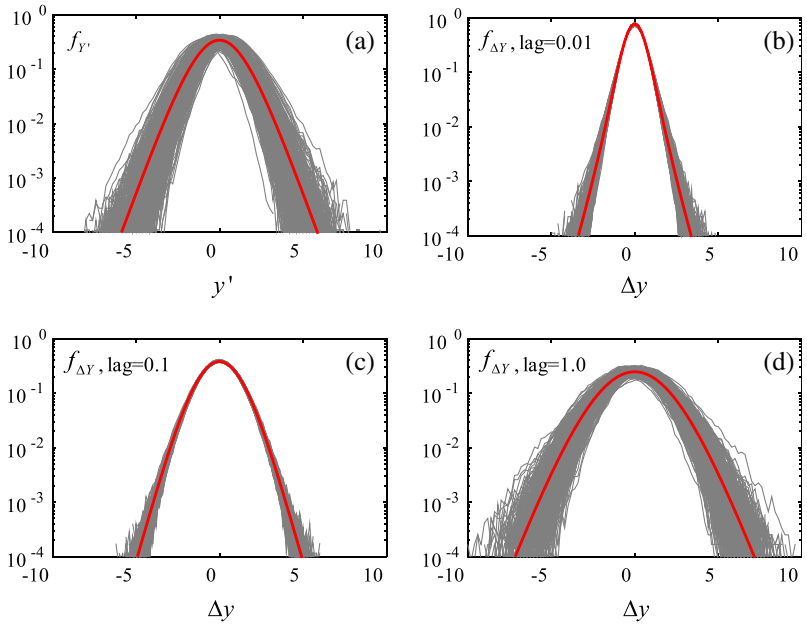

Fig. 3. Comparison of 500 selected spatial frequency distributions (gray curves) of (a) Y' and of (b, c, d) ΔY at three lags parallel to x_2 with corresponding theoretical pdfs (red curves) for TC1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$i = 1 \ldots N_{MC}$, with their ensemble counterparts and investigate the feasibility of estimating pdf parameters α, σ_G and ρ_G of the random functions Y' and ΔY based on such spatial samples. We start by computing the spatial mean M^{Y}_{1i} and variance M^{Y}_{2i} of Y' across each i-th realization and comparing it with $\langle Y' \rangle = 0$ and $\langle Y'^2 \rangle$. Table 1 lists minima, maxima, mean values and variances (var) of M^{Y}_{1i} and M^{Y}_{2i} across all realizations in each test case. The mean value of M^{Y}_{1i} is close to zero (on the order of $10^{-4} - 10^{-2}$) and that of M^{Y}_{2i} close to $\langle Y'^2 \rangle$ as given by (3) in all test cases. Though the minima and maxima of M^{Y}_{2i} differ measurably, its coefficients of variation (not listed) do not exceed 14%. Fig. 3 illustrates the correspondence between 500 selected spatial frequency distributions of Y' and of ΔY at three lags parallel to x_2, and their respective theoretical (ensemble) pdfs (2) and (5), for TC1. Results of similar quality were obtained for all other test cases (not shown).

We continue by computing $i = 1 \ldots N_{MC}$ estimates $\tilde{}\alpha_i$ of α and $\tilde{}\sigma_{Gi}$ of σ_G upon replacing the ensemble variance $\langle Y'^2 \rangle$ and kurtosis $\langle Y'^4 \rangle$ of Y' in (3) and (4) by their respective spatial counterparts, M^{Y}_{2i} and M^{Y}_{4i}. Figs. 4 and 5 depict histograms of these two sets of estimates, respectively, listing their mean, variance and coefficient of varia-tion for each test case. Mean values of these estimates are seen to represent the input parameters α and σ_G quite accurately with coefficients of variation on the order of 10^{-2} in all test cases. As

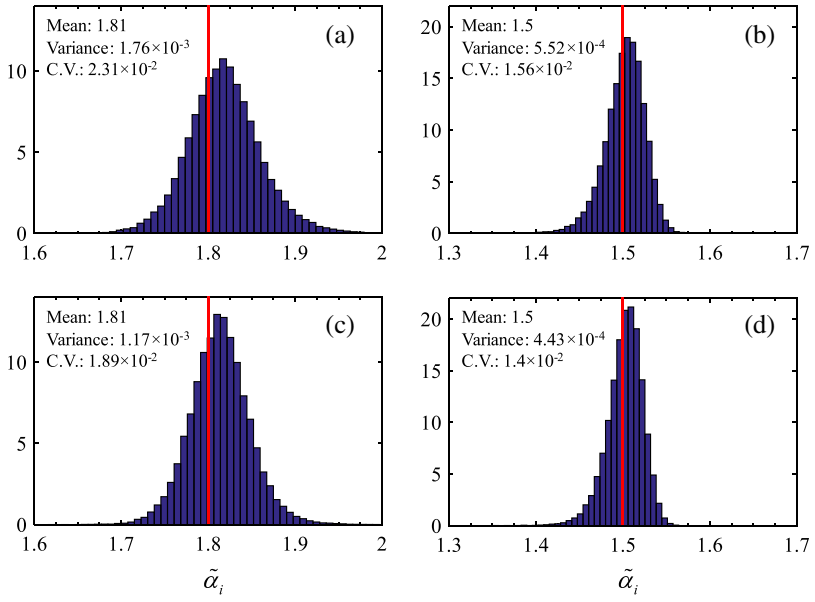

Fig. 4. Histograms of estimates $\tilde{\alpha}_i$ with corresponding mean, variance and coefficient of variation (C.V.) for (a) TC1, (b) TC2, (c) TC3, (d) TC4. Vertical red lines indicate input values α. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

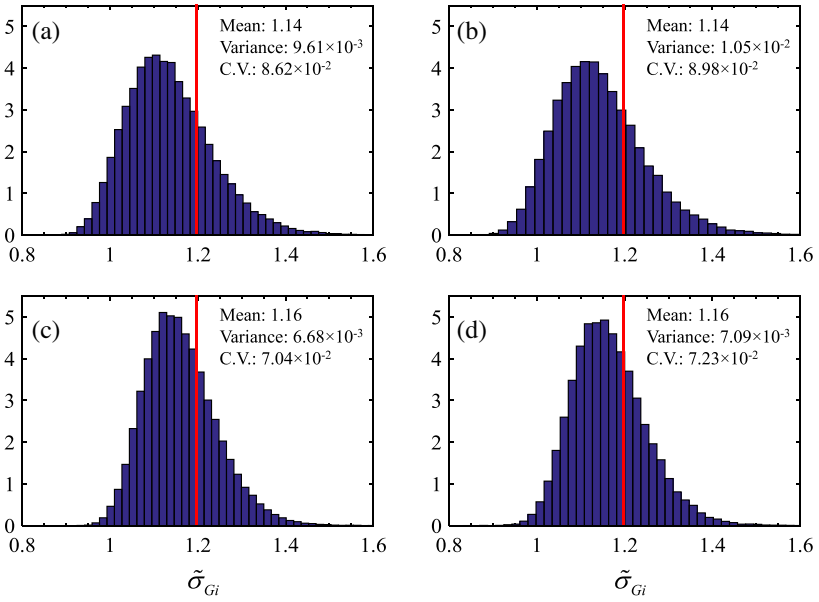

Fig. 5. Histograms of estimates $\tilde{\sigma}_{Gi}$ with corresponding mean, variance and coefficient of variation (C.V.) for (a) TC1, (b) TC2, (c) TC3, (d) TC4. Vertical red lines indicate input values σ_G. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

noted, however, by Riva et al. (2015) sampling Y' data without sampling their increments does not allow estimating the parameter ρ_G required to fully characterize the pdf of ΔY. Estimating all three parameters α, σ_G and ρ_G requires sampling both Y' and ΔY for a sufficiently wide range of lags. Doing so allows replacing $\langle Y'^2\rangle$, $\langle \Delta Y^2\rangle$ and $\langle \Delta Y^4\rangle$ in (3), (6) and (7) by their sample counterparts M_{2i}^{Y}, $M_{2i}^{\Delta Y}$ and $M_{4i}^{\Delta Y}$, respectively, to yield estimates $\hat{\alpha}_i$, $\hat{\sigma}_{Gi}$ and $\hat{\rho}_{Gi}$ for each realization $i = 1 \ldots N_{\mathrm{MC}}$. In contrast to α, σ_G and $\tilde{\alpha}_i$, $\tilde{\sigma}_{Gi}$ which are constant, $\hat{\alpha}_i$ and $\hat{\sigma}_{Gi}$ depend on ΔY and thus on lag as do ρ_G in (6) and (7) and its estimates $\hat{\rho}_{Gi}$. Fig. 6 compares mean of estimates $\hat{\alpha}_i$ and associated error bounds corresponding to their 10th and 90th percentile with mean of $\tilde{\alpha}_i$ and its error bounds as well as input values of α as functions of s_{x_1} for all test cases; Fig. 7 does the same for estimates of σ_G. Fig. 6 indicates that mean $\hat{\alpha}_i$ and $\tilde{\alpha}_i$ are generally close to each other and to α, the first having narrower error bounds at small and intermediate lags in two out of the four test cases. In Fig. 7 mean $\hat{\sigma}_{Gi}$ and $\tilde{\sigma}_{Gi}$ are generally closer to each other than to σ_G, having near identical error bounds at all lags in all four test cases. Results of comparable quality were obtained for increments parallel to x_2 (not shown).

Fig. 8 compares the input isotropic correlation function ρ_G with its mean estimates $\hat{\rho}_{Gi}$ and associated 10th as well as 90th

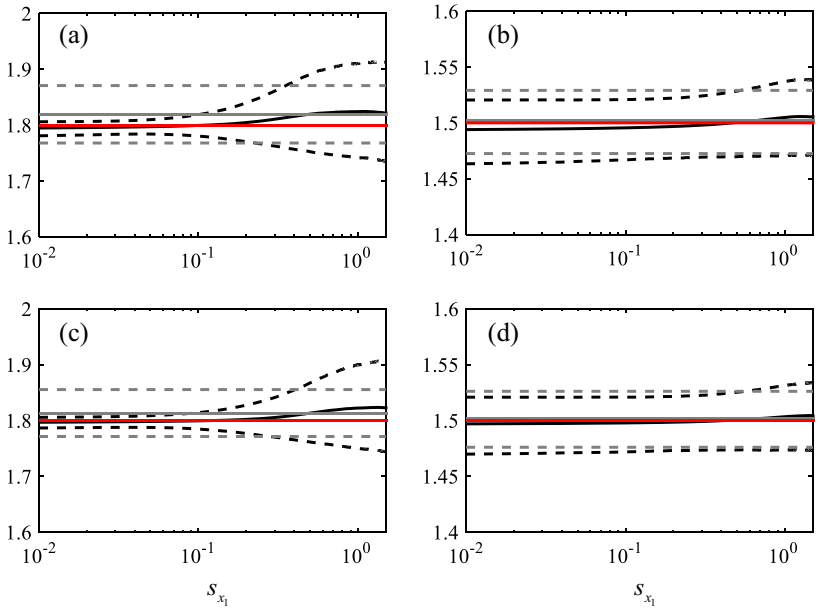

Fig. 6. Input parameters α (solid red), mean of estimates $\hat{\alpha}_i$ (solid black) and $\tilde{\alpha}_i$ (solid gray), error bounds corresponding to 10th and 90th percentile of $\hat{\alpha}_i$ (dashed black) and $\tilde{\alpha}_i$ (dashed gray) versus s_{x_1} for (a) TC1, (b) TC2, (c) TC3, (d) TC4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

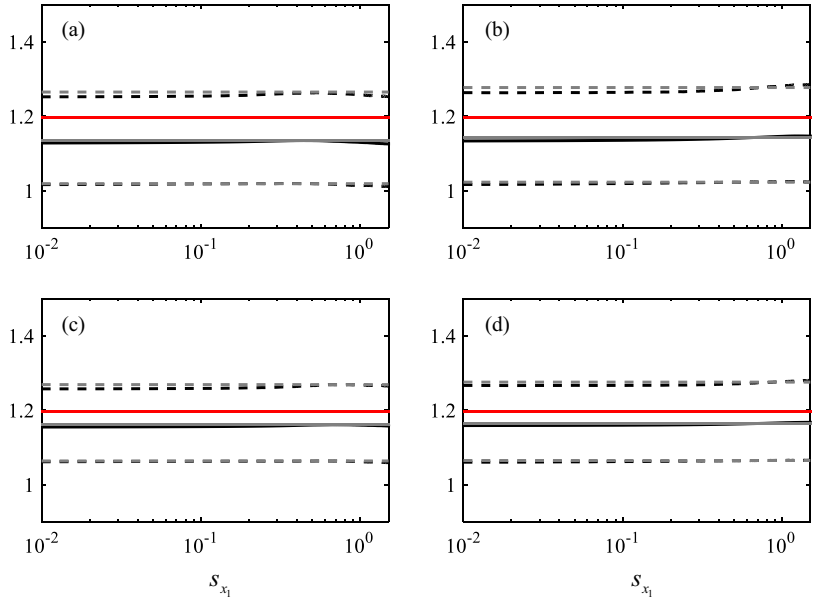

Fig. 7. Input parameter σ_G (solid red), mean of estimates $\hat{\sigma}_{Gi}$ (solid black) and $\tilde{\sigma}_{Gi}$ (solid gray), error bounds corresponding to 10th and 90th percentile of $\tilde{\sigma}_{Gi}$ (dashed black) and $\tilde{\sigma}_{Gi}$ (dashed gray) versus s_{x_1} for (a) TC1, (b) TC2, (c) TC3, (d) TC4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

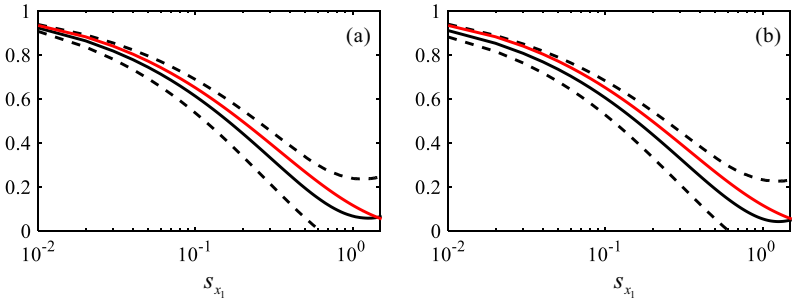

Fig. 8. Input isotropic correlation function ρ_G (solid red), mean of estimates $\hat{\rho}_{Gi}$ (solid black) and error bounds corresponding to 10th and 90th percentile of $\hat{\rho}_{Gi}$ (dashed black) versus s_{x_1} for (a) TC1, (b) TC2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

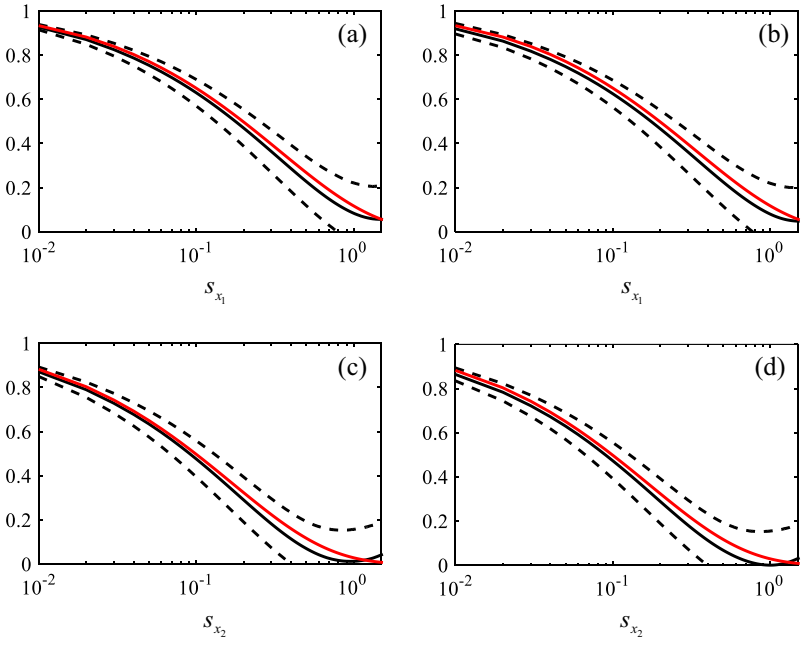

Fig. 9. Input anisotropic correlation functions ρ_G (solid red), mean of estimates $\hat{\rho}_{Gi}$ (solid black) and error bounds corresponding to 10th and 90th percentile of $\hat{\rho}_{Gi}$ (dashed black) versus s_{x_1} and s_{x_2} for (a–c) TC3 (b–d) TC4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

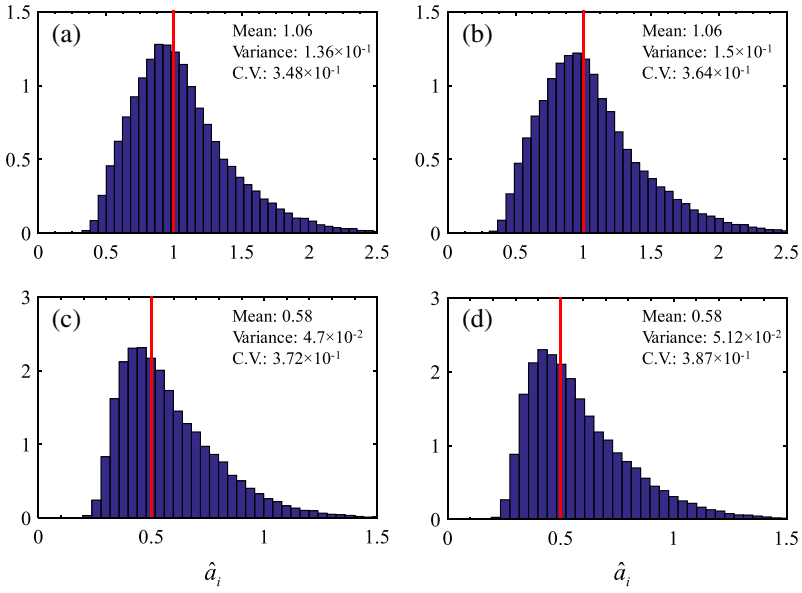

Fig. 10. Histograms of estimates $\hat{a}_i$ of a_i with corresponding mean, variance and coefficient of variation (C.V.) for (a) TC1, (b) TC2, (c) TC3, (d) TC4. Vertical red lines indicate input values a. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

percentile error bounds, plotted versus s_{x_1}, for TC1 and TC2; Fig. 9 presents similar plots versus s_{x_1} and s_{x_2} for test cases TC3 and TC4 in which ρ_G is anisotropic. In all four cases the accuracy of $\hat{\rho}_{Gi}$ improves and its error bounds narrow down with decreasing lag. Most important to note is that, in all cases, both isotropic and principal anisotropic ρ_G functions lie well within the error bounds of their estimates.

We end our analysis by computing estimates of directional integral scales $\hat{I}_{G_i x_1}$ and $\hat{I}_{G_i x_2}$, and of the anisotropy ratio $\hat{a}_i = \hat{I}_{G_i x_2} / \hat{I}_{G_i x_1}$, for each realization $i = 1 \ldots N_{MC}$ by integrating each estimate $\hat{\rho}_{Gi}$ along s_{x_1} and s_{x_2}, respectively. Fig. 10 plots histograms of $\hat{a}_i$ for all four test cases, lists their mean, variance and coefficient of variation, and indicates the corresponding input values of a. The estimates are seen to be accurate with coefficients of variation on order of 10^{-1}.

4. Conclusions

We presented and explored in two spatial dimensions a new multidimensional model of variables, Y, and increments, ΔY, that exhibit non-Gaussian statistical scaling. Our model, an extension of one proposed earlier by Riva et al. (2015), captures key aspects of such scaling, including statistical anisotropy, by treating Y as a generalized sub-Gaussian random function. The function is subordinated to truncated fractional Brownian motion (tfBm) defined by two cutoff scales, one proportional to a resolution measure of given data and the other to a measure of their sampling domain. Its subordinator is a non-negative random function (of space or time) all values of which are iid, that renders Y a scale mixture of Gaussian tfBm with random variances. Among various possible subordinators we selected in this paper one that is log-normally distributed. This renders Y Normal Log-Normal (NLN) with increments having a distribution that scales with separation distance or lag. This scaling, controlled by the correlation function of the underlying tfBm, causes the probability density function of ΔY to exhibit sharp peaks and heavy tails at small lags, both of which decay with increasing lag. Our new generalized sub-Gaussian model makes it easy to generate multidimensional, anisotropic random realiza-tions of Y and corresponding increments. Having done so 50,000 times on a square grid of 90,000 nodes using four separate sets of input parameters allowed us to demonstrate close correspon-dence (a) between key one- and two-point sample statistics of Y and ΔY at selected nodes and their theoretical (ensemble) counter-parts as well as (b) between these and corresponding spatial statis-tics across all nodes in individual realizations. This correspondence made it possible for us to demonstrate, with good results, a methodology to estimate all parameters characterizing the probability distributions of Y and ΔY on the basis of their values in individual realizations. Our ability to do so stems from the fact that, for given values of A and H in (11), the ratio I_Y/L between the integral scale of Y and the characteristic length scale L of the sampling domain is independent of L and tends to be a fraction of 1 (in other words, I_Y scales with L). The road is now open to consider subordinators other than log-normal, condition Y on measured values, interpolate and extrapolate its statistics between and outside of measurement points, and model flow as well as transport stochastically in randomly heterogeneous non-Gaussian environments.

Acknowledgements

Our work was supported in part through a contract between the University of Arizona and Vanderbilt University under the Consortium for Risk Evaluation with Stakeholder Participation (CRESP) III, funded by the U.S. Department of Energy. Funding from MIUR (Italian Ministry of Education, Universities and Research - PRIN2010-11; project: "Innovative methods for water resources under hydro-climatic uncertainty scenarios") is acknowledged.

References

Boffetta, G., Mazzino, A., Vulpiani, A., 2008. Twenty-five years of multifractals in fully developed turbulence. a tribute to Giovanni Paladin. J. Phys. A: Math. Theor. 41 (36), 363001. http://dx.doi.org/10.1088/1751-8113/41/36/363001.

Castaing, S., 1990. Velocity probability density functions of high Reynolds number turbulence. Phys. D: Nonlinear Phenom. 46 (2), 177–200. http://dx.doi.org/10.1016/0167-2789(90)90035-N.

Deutsch, C.V., Journel, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide, second ed. Oxford University Press, New York.

Di Federico, V., Neuman, S.P., Tartakovsky, D.M., 1999. Anisotropy, lacunarity, and upscaled conductivity and its autocovariance in multiscale random fields with truncated power variograms. Water Resour. Res. 35 (10), 2891–2908. http://dx.doi.org/10.1029/1999WR900158.

Ganti, V., Singh, A., Passalacqua, P., Foufoula-Georgiu, E., 2009. Subordinated Brownian motion model for sediment transport. Phys. Rev. E: Stat Nonlinear Soft Matter Phys. 80 (1), 011111. http://dx.doi.org/10.1103/PhysRevE.80.011111.

Guadagnini, A., Neuman, S.P., Riva, M., 2012. Numerical investigation of apparent multifractality of samples from processes subordinated to truncated fBm. Hydrol. Process. 26 (19), 2894–2908. http://dx.doi.org/10.1002/hyp.8358.

Guadagnini, A., Neuman, S.P., Schaap, M.G., Riva, M., 2013. Anisotropic statistical scaling of vadose zone hydraulic property estimates near Maricopa. Arizona. Water Resour. Res. 49 (12), 8463–8479. http://dx.doi.org/10.1002/2013WR014286.

Guadagnini, A., Neuman, S.P., Schaap, M.G., Riva, M., 2014. Anisotropic statistical scaling of soil and sediment texture in a stratified deep vadose zone near Maricopa, Arizona. Geoderma 214–215, 217–227. http://dx.doi.org/10.1016/j.geoderma.2013.09.008.

Guadagnini, A., Neuman, S.P., Nan, T., Riva, M., Winter, C.L., 2015. Scalable statistics of correlated random variables and extremes applied to deep borehole porosities. Hydrol. Earth Syst. Sci. 19 (2), 729–745. http://dx.doi.org/10.5194/hess-19-729-2015.

Kumar, P., Foufoula-Georgiou, E., 1993. A multicomponent decomposition of spatial rainfall fields: 2. Self-similarity in fluctuations. Water Resour. Res. 29 (8), 2533–2544. http://dx.doi.org/10.1029/93WR00549.

Liu, H.H., Molz, F.J., 1997. Comment on "Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations" by Scott Painter. Water Resour. Res. 33 (4), 907–908. http://dx.doi.org/10.1029/96WR03788.

Meerschaert, M.M., Kozubowski, T.J., Molz, F.J., Lu, S., 2004. Fractional Laplace model for hydraulic conductivity. Geophys. Res. Lett. 31 (8), L08501. http://dx.doi.org/10.1029/2003GL019320, 1–4..

Nan, T., Neuman, S. P., Riva, M., Guadagnini, A., 2015. Analyzing Randomly Fluctuating Hierarchical Variables and Extremes, Handbook of Groundwater Engineering, submitted.

Neuman, S.P., 2003. Relationship between juxtaposed, overlapping and fractal representations of multimodal spatial variability. Water Resour. Res. 9 (8), 1205. http://dx.doi.org/10.1029/2002WR001755.

Neuman, S.P., Guadagnini, A., Riva, M., Siena, M., 2013. Recent advances in statistical and scaling analysis of earth and environmental variables. In: Mishra, P.K., Kuhlman, K.L. (Eds.), Advances in Hydrogeology. Springer, New York, pp. 1–25.

Painter, S., 1996. Evidence for non-Gaussian scaling behavior of heterogeneous sedimentary formations. Water Resour. Res. 32 (5), 1183–1195. http://dx.doi.org/10.1029/96WR00286.

Painter, S., 2001. Flexible scaling model for use in random field simulation of hydraulic conductivity. Water Resour. Res. 37 (5), 1155–1163. http://dx.doi.org/10.1029/2000WR900394.

Riva, M., Neuman, S.P., Guadagnini, A., 2013a. Sub-Gaussian model of processes with heavy tailed distributions applied to air permeabilities of fractured tuff. Stoch. Environ. Res. Risk Assess. 27 (1), 195–207. http://dx.doi.org/10.1007/s00477-012-0576-y.

Riva, M., Neuman, S.P., Guadagnini, A., Siena, M., 2013b. Anisotropic scaling of Berea sandstone log air permeability statistics. Vadose Zone J. 12 (3). http://dx.doi.org/10.2136/vzj2012.0153.

Riva, M., Neuman, S.P., Guadagnini, A., 2015. New scaling model for variables and increments with heavy-tailed distributions. Water Resour. Res. 50. http://dx.doi.org/10.1002/2015WR016998.

Samorodnitsky, G., Taqqu, M.S., 1994. Stable Non-Gaussian Random Processes. Chapman & Hall, New York.

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. http://dx.doi.org/10.5194/hess-16-29-2012.

Siena, M., Guadagnini, A., Riva, M., Bijeljic, B., Pereira Nunes, J.P., Blunt, M.J., 2014. Statistical scaling of pore-scale Lagrangian velocities in natural porous media. Phys. Rev. E: Stat, Nonlinear, Soft Matter Phys. 90 (2), 023013. http://dx.doi.org/10.1103/PhysRevE.90.023013, 1–8..

von Papen, M., Saur, J., Alexandrova, O., 2014. Turbulent magnetic field fluctuations in Saturn's magnetosphere. J. Geophys. Res.: Space Phys. 119 (4), 2797–2818. http://dx.doi.org/10.1002/2013JA019542.

Yang, C.-Y., Hsu, K.-C., Chen, K.-C., 2009. The use of the Levy-stable distribution for geophysical data analysis. Hydrogeol. J. 17 (5), 1265–1273. http://dx.doi.org/10.1007/s10040-008-0411-1.