WE-NEED

WatEr NEEDs, availability, quality and sustainability

Deliverable Number:	D1.2
Work package number:	WP1
Deliverable title	REPORT ON NEW SUB-GAUSSIAN MODELS
Type	Research
Dissemination Level	Public
Lead participant	Politecnico di Milano
Contributing scientists and other personnel	Alberto Guadagnini, Monica Riva, Martina Siena
Scheduled delivery date	20 March 2017
Actual / forecast delivery date	22 March 2017

Deliverable summary.

This report presents a new statistical model (generalized sub-Gaussian, GSG) for the interpretation of hydrological properties, Y, as well as many other variable, exhibiting a clear non-Gaussian behavior. One common manifestation of non-Gaussianity is that whereas frequency distributions, pdfs, of Y often exhibit mild peaks and light tails, those of increments are generally symmetric with peaks that grow sharper, and tails that become heavier, as separation scale or lag between pairs of Y values decreases. We derive analytical expressions for pdfs of data and the associated spatial increments as well as corresponding lead statistical moments. In our GSG model the peak and tails of the increments pdf scale with lag, in line with the characteristic behavior exhibited by many hydrological variables. The model allows one to estimate, accurately and efficiently, all relevant parameters by analyzing jointly sample moments of data and incremental series. We illustrate key features of our new model and method of inference on a set of neutron porosity data from a deep borehole. Future developments of this work include the analysis of lead-order effects that non-Gaussian heterogeneity described by the GSG model have on the stochastic description of flow and transport.

D1.2

Report on new sub-Gaussian Models

Contents

1. Introduction

Efforts aimed at the characterization of the spatial variability of subsurface properties typically focus on the identification of a geostatistical model consistent with the key features exhibited by data and their main statistics. The study of a wide range of datasets of Earth, environmental and several other variables frequently highlights that increments $\Delta Y(s) = Y(\boldsymbol{x}) - Y(\boldsymbol{y})$ of a variable Y calculated between vector locations $\boldsymbol{x}$ and $\boldsymbol{y}$ ($s = ||\boldsymbol{x} - \boldsymbol{y}||$ being separation scale or lag) are characterized by sample distributions with sharp peaks and heavy tails, a behavior which tends to become increasingly marked as the lag decreases. Documented examples of such behavior for hydrological and soil science variables include, among others, datasets of porosity (Painter, 1996; Guadagnini et al., 2014, 2015; Riva et al., 2015), permeability (Painter, 1996; Siena et al., 2012; Riva et al., 2013a, b), hydraulic conductivity (Liu and Molz, 1997; Meerschaert et al., 2004; Guadagnini et al., 2013), hydraulic parameters characterizing unsaturated soils (Guadagnini et al., 2013), soil and sediment texture data (Guadagnini et al., 2014) and pore scale velocities (Siena et al., 2014). It is then clear that the assumption of Gaussianity for Y is not consistent with the above-described characteristics of statistical scaling displayed by the sample probability distribution (and main statistical moments) of increments. A statistical model that captures the scale-dependent features of the probability density of ΔY in a unified and consistent manner has recently been proposed by Riva et al. (2015) and Panzeri et al. (2016). These authors view $Y(\boldsymbol{x}) = \langle Y \rangle + Y'(\boldsymbol{x})$ as a spatial random field, $\langle Y \rangle$ and $Y'(\boldsymbol{x})$ respectively being the ensemble mean and a local zero-mean fluctuation. The latter can be expressed through the following generalized sub-Gaussian (GSG) model

$$Y'(\mathbf{x}) = U(\mathbf{x})G(\mathbf{x}) \qquad (1)$$

where $\boldsymbol{x}$ is a position vector, $G(\boldsymbol{x})$ is (generally, but not necessarily) a multi-scale Gaussian random field and $U(\boldsymbol{x})$ is a subordinator independent of G. The subordinator U consists of statistically independent identically distributed (iid) non-negative random values at all points $\boldsymbol{x}$**.** Several choices about the probability density function shape, pdf, of U are possible. In this context, here we study the main features of the GSG model (1) by considering (*i*) log-normal and (*ii*) Gamma distributional form of the subordinator U . The remainder of the study is structured as follows. Section 2 illustrates the details of the analytical formulation of the GSG

model. Section 3 describes methods of inference associated with the diverse subordinators. In Section 4 we present an application of the GSG models to field data and we demonstrate the remarkable features of our GSG model to interpret within a unique theoretical framework the scaling properties of sample pdf of Y and of its of increments.

2. Theoretical framework

We introduce the following notation to define Y' at two points, $\mathbf{x}_1$ and $\mathbf{x}_2$,

$$Y'(\mathbf{x}_1) = U(\mathbf{x}_1)G(\mathbf{x}_1) = Y_1 = U_1 G_1, \qquad Y'(\mathbf{x}_2) = U(\mathbf{x}_2)G(\mathbf{x}_2) = Y_2 = U_2 G_2 . \tag{2}$$

The bivariate pdf of Y_1 and Y_2 is

$$f_{Y_1,Y_2}(y_1,y_2) = \int_0^\infty\int_0^\infty f_{U_1}(u_1)\, f_{U_2}(u_2)\, f_{G_1G_2}\left(\frac{y_1}{u_1},\frac{y_2}{u_2}\right)\frac{du_2}{u_2}\,\frac{du_1}{u_1} \tag{3}$$

where $f_{U_i}(u_i)$ is the pdf of U_i (i = 1, 2) and $f_{G_1G_2}$ is the bivariate pdf of (G_1,G_2) given by

$$f_{G_1G_2}\left(\frac{y_1}{u_1},\frac{y_2}{u_2}\right) = \frac{e^{-\frac{1}{2\sigma_G^2(1-\rho_G^2)}\left(\frac{y_1^2}{u_1^2}+\frac{y_2^2}{u_2^2}-2\rho_G\frac{y_1}{u_1}\frac{y_2}{u_2}\right)}}{2\pi\sigma_G^2\sqrt{1-\rho_G^2}} . \tag{4}$$

Here, σ_G^2 is the variance of G and ρ_G is the coefficient of correlation between G_1 and G_2. Substituting (4) into (3) yields the following bivariate pdf of Y_1 and Y_2,

$$f_{Y_1,Y_2}(y_1,y_2) = \frac{1}{2\pi\sigma_G^2\sqrt{1-\rho_G^2}}\int_0^\infty\int_0^\infty f_{U_1}(u_1)\, f_{U_2}(u_2)\, e^{-\frac{1}{2\sigma_G^2(1-\rho_G^2)}\left(\frac{y_1^2}{u_1^2}+\frac{y_2^2}{u_2^2}-2\rho_G\frac{y_1}{u_1}\frac{y_2}{u_2}\right)}\frac{du_2}{u_2}\,\frac{du_1}{u_1} . \tag{5}$$

The marginal pdf of Y' can been obtained as

$$\begin{aligned} f_{Y'}(y) &= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1,y_2)\,dy_1 \\ &= \frac{1}{2\pi\sigma_G^2\sqrt{1-\rho_G^2}}\int_{u_1=0}^{\infty}\int_{u_2=0}^{\infty} f_{U_1}(u_1)\, f_{U_2}(u_2)\left\{\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_G^2(1-\rho_G^2)}\left[\frac{y_1^2}{u_1^2}+\frac{y_2^2}{u_2^2}-2\rho_G\frac{y_1}{u_1}\frac{y_2}{u_2}\right]}dy_1\right\}\frac{du_2}{u_2}\,\frac{du_1}{u_1} = \\ &= \frac{1}{\sqrt{2\pi}\sigma_G}\int_0^\infty\int_0^\infty f_{U_1}(u_1)\, f_{U_2}(u_2)\, e^{-\frac{1}{2\sigma_G^2}\frac{y^2}{u_2^2}}\frac{du_2}{u_2}du_1 = \frac{1}{\sqrt{2\pi}\sigma_G}\int_0^\infty f_{U_2}(u_2)\, e^{-\frac{1}{2\sigma_G^2}\frac{y^2}{u_2^2}}\frac{du_2}{u_2} \end{aligned} \tag{6}$$

The first order statistical moment of Y' is identically zero, whereas variance, kurtosis and q-th order moment can be expressed respectively as

$$\left\langle Y'^2\right\rangle = \int_{-\infty}^{+\infty} y^2 f_{Y'}(y)dy = \sigma_G^2\int_0^\infty u_2^{\,2} f_{U_2}(u_2)\,du_2 = \sigma_G^2\left\langle U^2\right\rangle, \tag{7}$$

$$\left\langle Y'^4 \right\rangle = \int_{-\infty}^{+\infty} y^4 f_{Y'}(y)\,dy = 3\sigma_G^4 \int_0^{\infty} u_2{}^4 f_{U_2}(u_2)\,du_2 = 3\sigma_G^4 \left\langle U^4 \right\rangle, \tag{8}$$

while the q-th order moment is

$$\left\langle Y'^q \right\rangle = \int_{-\infty}^{+\infty} y^q f_Y(y)dy = \frac{2^{\frac{q-1}{2}}\left(1+(-1)^q\right)}{\sqrt{2\pi}}\Gamma\left(\frac{1+q}{2}\right)\sigma_G^q \int_0^{\infty} u_2{}^q f_{U_2}(u_2)\,du_2, \text{ with } q > 1. \tag{9}$$

The probability density function of incremental data, ΔY, is given by

$$\begin{aligned} f_{\Delta Y}(\Delta y) &= \int_{-\infty}^{\infty} f_{Y_1Y_2}(\Delta y + y_2, y_2)\,dy_2 \\ &= \frac{1}{2\pi\sigma_G^2\sqrt{1-\rho_G^2}} \int_0^{\infty}\int_0^{\infty} f_{U_1}(u_1) f_{U_2}(u_2) \left\{ \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_G^2\left(1-\rho_G^2\right)}\left[\frac{(\Delta y+y_2)^2}{u_1{}^2}+\frac{y_2{}^2}{u_2{}^2}-2\rho_G\frac{\Delta y+y_2}{u_1}\frac{y_2}{u_2}\right]} dy_2 \right\} \frac{du_2}{u_2}\frac{du_1}{u_1} \\ &= \frac{1}{\sqrt{2\pi}\sigma_G} \int_0^{\infty}\int_0^{\infty} \frac{f_{U_1}(u_1) f_{U_2}(u_2)}{\sqrt{\left(u_1{}^2+u_2{}^2-2\rho_G u_1 u_2\right)}} e^{-\frac{1}{\sigma_G^2\left(u_1{}^2+u_2{}^2-2\rho_G u_1 u_2\right)}\frac{\Delta y^2}{2}} du_2 du_1 \end{aligned} \tag{10}$$

The first order moment of ΔY is identically zero, whereas variance, kurtosis and q-th order moment can be expressed respectively as

$$\left\langle \Delta Y^2 \right\rangle = \int_{-\infty}^{+\infty} \Delta y^2 f_{\Delta Y}(\Delta y)d(\Delta y) = \sigma_G^2 \int_0^{\infty}\int_0^{\infty} \left(u_1{}^2+u_2{}^2-2\rho_G u_1 u_2\right) f_{U_1}(u_1) f_{U_2}(u_2)\; du_2 du_1 \tag{11}$$

$$\left\langle \Delta Y^4 \right\rangle = \int_{-\infty}^{+\infty} \Delta y^4 f_{\Delta Y}(\Delta y)d(\Delta y) = 3\sigma_G^4 \int_0^{\infty}\int_0^{\infty} \left(u_1{}^2+u_2{}^2-2\rho_G u_1 u_2\right)^2 f_{U_1}(u_1) f_{U_2}(u_2)\; du_2 du_1 \tag{12}$$

$$\begin{aligned} \left\langle \Delta Y^q \right\rangle &= \int_{-\infty}^{+\infty} \Delta y^q f_{\Delta Y}(\Delta y)d(\Delta y) = \\ &= \frac{\sigma_G^q}{\sqrt{2\pi}} 2^{\frac{q-1}{2}}\left(1+(-1)^q\right)\Gamma\left(\frac{q+1}{2}\right) \int_0^{\infty}\int_0^{\infty} \left(u_1{}^2+u_2{}^2-2\rho_G u_1 u_2\right)^{\frac{q}{2}} f_{U_1}(u_1) f_{U_2}(u_2) du_2 du_1 \end{aligned} \tag{13}$$

while the q-th order moment of the absolute value of increments is given by

$$\begin{aligned} \left\langle |\Delta Y|^q \right\rangle &= \int_{-\infty}^{+\infty} |\Delta y|^q f_{\Delta Y}(\Delta y)\,d(\Delta y) = \\ &= \frac{\sigma_G^q}{\sqrt{2\pi}} 2^{\frac{q+1}{2}}\Gamma\left(\frac{q+1}{2}\right) \int_0^{\infty}\int_0^{\infty} \sqrt{\left(u_1{}^2+u_2{}^2-2\rho_G u_1 u_2\right)^q} f_{U_1}(u_1) f_{U_2}(u_2)\; du_2 du_1 \end{aligned} \tag{14}$$

The Covariance of Y between two points x_1 and x_2, is defined as:

$$C_Y(x_1,x_2)=\langle Y'(\boldsymbol{x}_1)Y'(\boldsymbol{x}_2)\rangle \\ =\langle U(\boldsymbol{x}_1)U(\boldsymbol{x}_2)\rangle\langle G(\boldsymbol{x}_2)G(\boldsymbol{x}_1)\rangle=\langle U(\boldsymbol{x}_1)U(\boldsymbol{x}_2)\rangle C_G(\boldsymbol{x}_1,\boldsymbol{x}_2) \tag{15}$$

where $C_G=\sigma_G^2\rho_G$ is the covariance of G. From (15) one derives

$$C_Y(0)=\sigma_Y^2=\langle U^2\rangle\sigma_G^2,\quad C_Y(s>0)=\langle U\rangle^2 C_G(s),\qquad \rho_Y=\frac{C_Y(s)}{\sigma_Y^2}=\frac{\langle U\rangle^2}{\langle U^2\rangle}\rho_G(s) \tag{16}$$

Note that according to (16) the sub-Gaussian covariance C_Y is discontinuous at r = 0, exhibiting a nugget effect. This observation also implies that nugget effects, attributed in the literature to variability of Y at scales smaller than the sampling interval and/or to measurement errors, may in fact be (at least in part) a symptom of non-Gaussianity.

2.1 Log-normal subordinator

Here, we assume that U_1 and U_2 are lognormally distributed according to $\ln N\left(0,(2-\alpha)^2\right)$, i.e.

$$f_{U_i}(u_i)=\frac{e^{-\frac{\ln^2 u_i}{2(2-\alpha)^2}}}{\sqrt{2\pi}u_i(2-\alpha)} \qquad \text{with } i=1,2 \quad \text{and } \alpha<2. \tag{17}$$

Substituting (17) into (6) yield the following marginal pdf of Y'

$$f_{Y'}(y)=\frac{1}{2\pi\sigma_G(2-\alpha)}\int_0^\infty e^{-\frac{1}{2}\left[\frac{\ln^2 u}{(2-\alpha)^2}+\frac{y^2}{\sigma_G^2u^2}\right]}\frac{du}{u^2}. \tag{18}$$

Equation (18) coincides with (10) of Riva et al. (2015). Setting $x/\sigma_G = u$, (18) becomes

$$f_{Y'}(y)=\frac{1}{2\pi(2-\alpha)}\int_0^\infty e^{-\frac{1}{2}\left(\frac{\ln^2 x/\sigma_G}{(2-\alpha)^2}+\frac{y^2}{x^2}\right)}\frac{dx}{x^2}. \tag{19}$$

Note that (19) coincides with a Normal-Log-Normal distribution, NLN. The latter have been shown to well represent some financial (Clark, 1973) and environmental (Guadagnini et al.. 2015) data. Making use of (7) - (9), variance, kurtosis, standardized kurtosis and q-th order moments of Y' are respectively given by

$$\langle Y'^2\rangle=\sigma_G^2e^{2(2-\alpha)^2} \tag{20}$$

$$\langle Y'^4\rangle=3\sigma_G^4e^{8(2-\alpha)^2} \tag{21a}$$

$$\kappa_Y = \frac{\langle Y'^4 \rangle}{\langle Y'^2 \rangle^2} = 3e^{4(2-\alpha)^2} \tag{21b}$$

$$\langle Y'^q \rangle = \frac{2^{\frac{q-1}{2}}\left(1+(-1)^q\right)}{\sqrt{2\pi}}\Gamma\left(\frac{1+q}{2}\right)\sigma_G^q e^{(2-\alpha)^2\frac{q^2}{2}} \quad \text{with } q > 1. \tag{22}$$

Substituting (17) into (10) yields the following pdf of increments ΔY

$$f_{\Delta Y}(\Delta y) = \frac{1}{2\pi^2(2-\alpha)^2\sigma_G}\sqrt{\frac{\pi}{2}}\int_0^\infty\int_0^\infty \frac{e^{-\frac{1}{2}\left[\frac{1}{(2-\alpha)^2}\left(\ln^2 u_2+\ln^2 u_1\right)+\frac{\Delta y^2}{\sigma_G^2\left(u_1^2+u_2^2-2\rho_G u_1 u_2\right)}\right]}}{u_2 u_1\sqrt{u_1^2+u_2^2-2\rho_G u_1 u_2}}du_2 du_1 \tag{23}$$

The latter coincides with (17) of Riva et al. (2015). Making use of (11)-(12), variance, kurtosis and standardized kurtosis of ΔY are then given, respectively, by

$$\langle \Delta Y^2 \rangle = 2\sigma_G^2 e^{(2-\alpha)^2}\left(e^{(2-\alpha)^2}-\rho_G\right) \tag{24}$$

$$\langle \Delta Y^4 \rangle = 6\sigma_G^4 e^{4(2-\alpha)^2}\left(1+2\rho_G^2-4\rho_G e^{(2-\alpha)^2}+e^{4(2-\alpha)^2}\right) \tag{25a}$$

$$\kappa_{\Delta Y} = \frac{\langle \Delta Y^4 \rangle}{\langle \Delta Y^2 \rangle^2} = 3e^{2(2-\alpha)^2}\left\{1+\frac{1}{2}\left(\frac{e^{2(2-\alpha)^2}-1}{e^{(2-\alpha)^2}-\rho_G}\right)^2\right\}. \tag{25b}$$

As $\alpha \to 2$, distributions of Y' and ΔY tends to the Gaussian. Otherwise the shape of $f_{\Delta Y}$ scales with the correlation coefficient of G or, equivalently, with lag as quantified by $\kappa_{\Delta Y}$. In other words, the peak of $f_{\Delta Y}$ sharpens and its tails become heavier with ρ_G. Figure 1 illustrates how excess kurtosis, $\kappa_{\Delta Y}-3$, varies with ρ_G and α. At small lags (large ρ_G) $\kappa_{\Delta Y}-3$ exceeds zero by a significant margin, even at large values of α (when the pdf of Y' is near-Gaussian). Excess kurtosis decreases as ρ_G decreases (lag increases), rendering the peak of $f_{\Delta Y}$ less sharp and its tails lighter. When $\alpha \geq 1.8$, the asymptotic value of $\kappa_{\Delta Y}-3$ at large lags is very small (<< 1) and $f_{\Delta Y}$ is virtually Gaussian. Included in Figure 1 are horizontal lines depicting excess kurtosis of the pdf of Y', $\kappa_Y - 3$. From (21b) and (25b) it follows that when $\alpha > 2-\sqrt{\ln 3} \approx 0.95$, $f_{\Delta Y}$ at small lags has sharper peaks and heavier tails than does $f_{Y'}$, the opposite being true at large lags. When $\alpha < 2-\sqrt{\ln 3}$ the pdf of Y' has higher peaks and

heavier tails than the pdfs of ΔY regardless of lag. This behavior of $f_{\Delta Y}$ is indicated by Figure 1. Figure 2 depicts on arithmetic and semi logarithmic scales $f_{\Delta Y}$ for $\sigma_G = 1.0$, $\alpha = 1.7$ and three values of ρ_G. Also shown for comparison is a Gaussian distribution having the same mean and variance as ΔY.

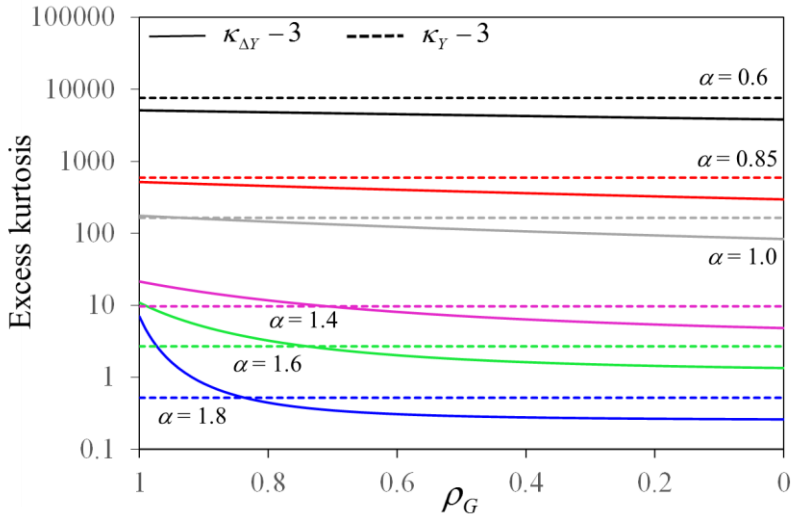

Figure 1. Excess kurtosis of ΔY (continuous curves) and of Y' (horizontal dashed lines) versus ρ_G for five values of α.

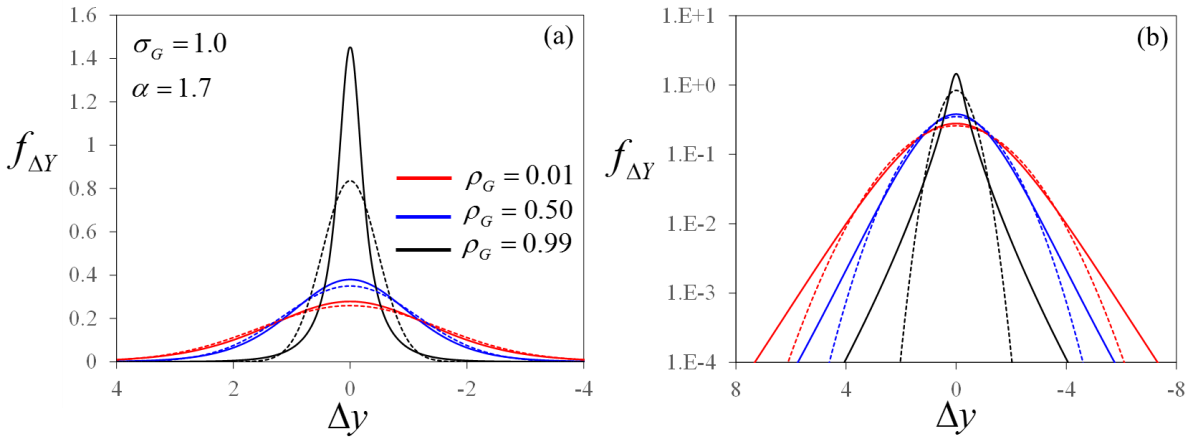

Figure 2. $f_{\Delta Y}$ (23) on arithmetic (a) and semi-logarithmic (b) scales for $\sigma_G = 1.0$, $\alpha = 1.7$ and three values of ρ_G (continuous curves). Also shown are Gaussian distributions having the same mean and variance as ΔY (dashed curves).

As noted earlier, $f_{\Delta Y}$ exhibits sharp peak and heavy tails when ρ_G = 0.99 (at small lag) and becomes virtually Gaussian as ρ_G decreases (lag increases).

The *q*-th order moments of ΔY and $|\Delta Y|$ are given respectively by

$$\left\langle \Delta Y^q \right\rangle = \frac{\sigma_G^q}{\sqrt{2\pi}} 2^{\frac{q-1}{2}} \left(1+(-1)^q\right) \Gamma\left(\frac{q+1}{2}\right) \frac{1}{2\pi(2-\alpha)^2} \int_0^\infty \int_0^\infty \left(u_1^2 + u_2^2 - 2\rho_G u_1 u_2\right)^{\frac{q}{2}} e^{-\frac{\ln^2 u_2 + \ln^2 u_1}{2(2-\alpha)^2}} \frac{du_2}{u_2} \frac{du_1}{u_1} \tag{26}$$

$$\left\langle |\Delta Y|^q \right\rangle = \frac{\sigma_G^q}{\sqrt{\pi}} \frac{2^{\frac{q}{2}}}{2\pi(2-\alpha)^2} \Gamma\left(\frac{q+1}{2}\right) \int_0^\infty \int_0^\infty \left(u_1^2 + u_2^2 - 2\rho_G u_1 u_2\right)^{\frac{q}{2}} e^{-\frac{\ln^2 u_2 + \ln^2 u_1}{2(2-\alpha)^2}} \frac{du_2}{u_2} \frac{du_1}{u_1}. \tag{27}$$

The variogram of *Y'*, γ_Y, can be derived from (24) as

$$\gamma_Y = \frac{\left\langle \Delta Y^2 \right\rangle}{2} = \sigma_G^2 e^{(2-\alpha)^2} \left(e^{(2-\alpha)^2} - \rho_G\right) = \sigma_G^2 e^{(2-\alpha)^2} \left(e^{(2-\alpha)^2} - 1\right) + \gamma_G e^{(2-\alpha)^2}, \tag{28}$$

γ_G being the variogram of G. The covariance, C_Y, is in turn

$$C_Y(0) = e^{2(2-\alpha)^2} \sigma_G^2, \quad C_Y(s>0) = e^{(2-\alpha)^2} C_G, \tag{29}$$

Equation (29) in turn yields the integral scale of *Y'* to be

$$I_Y = e^{-(2-\alpha)^2} I_G, \tag{30}$$

being I_G the integral scale of G. It is thus seen that a lognormal subordinator dampens, but does not destroy, the covariance structure of *G*; the smaller is α the shorter is the integral scale of *Y'*. When $\alpha \to 2$, $I_Y \to I_G$. The integral scale of *Y'* vanishes only in the limit as $\alpha \to -\infty$, i.e. when the variance of subordinator *U* tends to infinity.

2.2 Gamma subordinator

Here, we assume that U_1 and U_2 are distributed according to a Gamma distribution, i.e.,

$$f_{U_i}(u_i) = \frac{u_i^{k-1} e^{-\frac{u_i}{\theta}}}{\Gamma(k)\theta^k}, \qquad \text{with } i = 1, 2;\ k > 0;\ \theta > 0, \tag{31a}$$

where *k* and θ respectively are the shape and scale parameter and

$$\Gamma(k) = \int_0^\infty x^{k-1} e^{-x}\, dx \qquad k > 0 \tag{31b}$$

is the Gamma function. Note that the Exponential subordinator can be obtained by (31a) setting $k = 1$ and $\lambda = 1/\theta$. According to (31a) mean and variance of U_i are given respectively by $k\theta$ and $k\theta^2$.

Substituting (31a) into (6) yield the following marginal pdf of Y'

$$f_{Y'}(y) = \frac{1}{\sqrt{2\pi}\sigma_G}\frac{1}{\Gamma(k)\theta^k}\int_0^\infty u_2^{k-2} e^{-\frac{u_2}{\theta}} \exp\left[-\frac{1}{2\sigma_G^2}\frac{y^2}{u_2^2}\right] du_2 . \tag{32}$$

Making use of (7) - (9) and (31a), variance, kurtosis, standardized kurtosis and q-th order moments of Y' are respectively given by

$$\left\langle Y'^2 \right\rangle = \sigma_G^2\theta^2 \frac{\Gamma(2+k)}{\Gamma(k)} = \sigma_G^2\theta^2 k(1+k) \tag{33}$$

$$\left\langle Y'^4 \right\rangle = 3\sigma_G^4\theta^4 \frac{\Gamma(4+k)}{\Gamma(k)} = 3\sigma_G^4\theta^4 k(k+1)(k+2)(k+3) \tag{34a}$$

$$\kappa_Y = \frac{\left\langle Y'^4 \right\rangle}{\left\langle Y'^2 \right\rangle^2} = 3\frac{(k+2)(k+3)}{k(1+k)} \tag{34b}$$

$$\left\langle Y'^q \right\rangle = \frac{2^{\frac{q-1}{2}}\left(1+(-1)^q\right)}{\sqrt{2\pi}}\Gamma\left(\frac{1+q}{2}\right)\sigma_G^q \frac{\Gamma(k+q)}{\Gamma(k)}\theta^q . \tag{35}$$

Figure 3 depicts on semi logarithmic scales $f_{Y'}(y)$ for $\sigma_G = 1.0$, $\theta = 1$ and five values of k. Also shown for comparison is a Gaussian distribution having the same mean and variance as Y'. In Figure 4 we investigate combinations of parameters (k, θ, σ_G^2) leading to pdfs of Y' that are consistent with typical permeability field data. Figure 4a depicts the variance Y', as rendered by (33) versus k for two selected values of θ, i.e., $\theta = 0.2, 0.5$. Figure 4b depicts the pdf of $f_{U_i}(u_i)$ (31a), for $k = 10$ and $\theta = 0.2, 0.5$. Figure 4c reports $f_{Y'}(y)$ for $\sigma_G^2 = 1$, $k = 10$ and $\theta = 0.2, 0.5$.

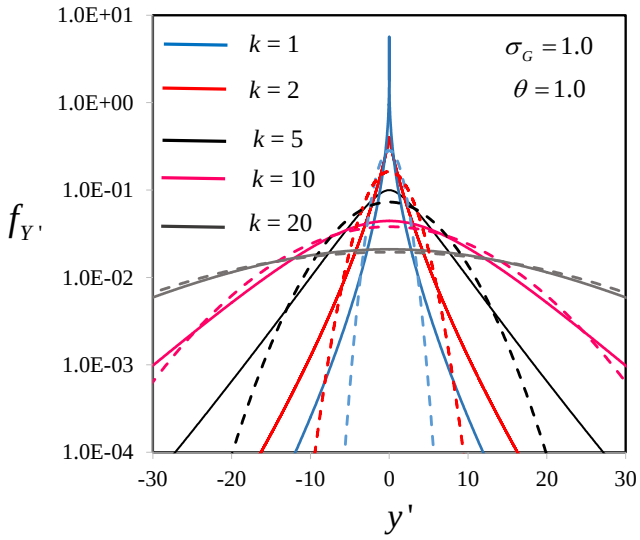

Figure 3. $f_{Y'}$ (32) on semi-logarithmic scales for $\sigma_G^2 = 1.0$, $\theta = 1$ and five values of k Also shown are Gaussian distributions having the same mean and variance as Y' (dashed curves).

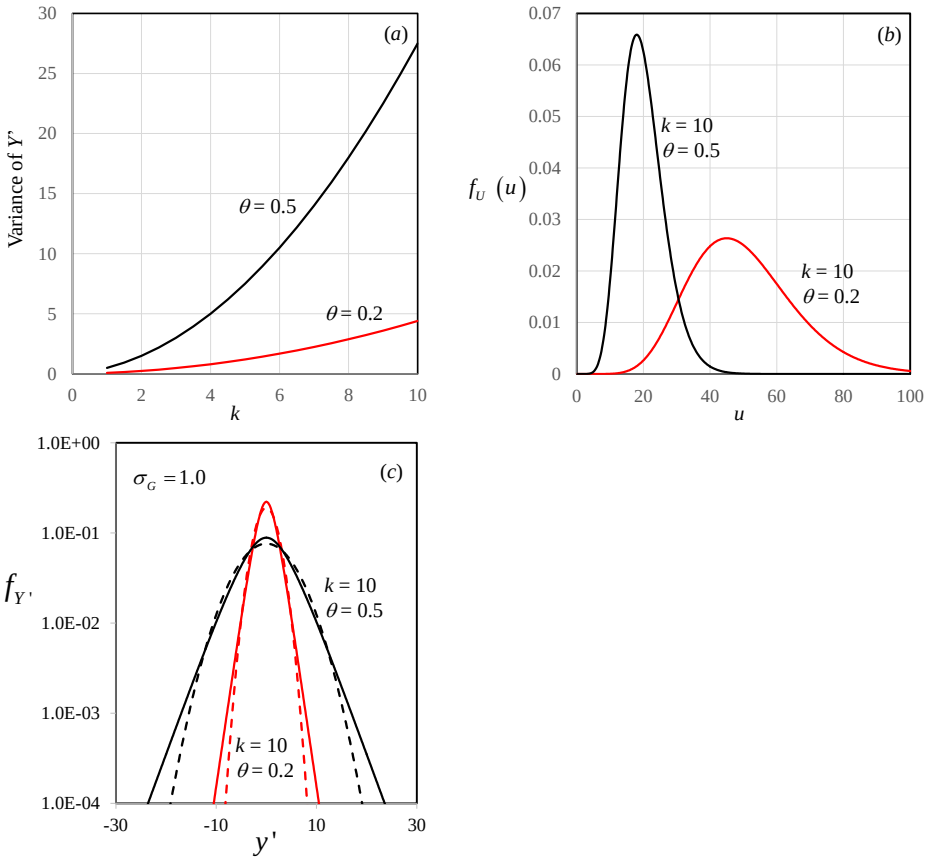

Figure 4. (*a*) Variance of Y' (33) versus k for $\theta = 0.2, 0.5$; (*b*) $f_{U_i}(u_i)$ (31a) for $k = 10$ and $\theta = 0.2, 0.5$; (*c*) $f_{Y'}$ (32) for $\sigma_G^2 = 1.0$, $k = 10$ and $\theta = 0.2, 0.5$. Also shown are Gaussian distributions having the same mean and variance as Y' (dashed curves).

Substituting (31a) into (10) yields the following pdf of ΔY

$$f_{\Delta Y}(\Delta y)=\frac{1}{\sqrt{2\pi}\sigma_G\Gamma^2(k)\theta^{2k}}\int_0^\infty\int_0^\infty \frac{(u_1u_2)^{k-1}e^{-\frac{1}{\theta}(u_1+u_2)-\frac{1}{\sigma_G^2\left(u_1^2+u_2^2-2\rho_G u_1u_2\right)}\frac{\Delta y^2}{2}}}{\sqrt{\left(u_1^2+u_2^2-2\rho_G u_1u_2\right)}}\,du_2du_1. \tag{36}$$

Making use of (11)-(12), variance, kurtosis and standardized kurtosis of ΔY are then given, respectively, by

$$\left\langle \Delta Y^2\right\rangle = 2\sigma_G^2k\theta^2\left(1+\left(1-\rho_G\right)k\right) \tag{37}$$

$$\left\langle \Delta Y^4\right\rangle = 12\sigma_G^4k(1+k)\theta^4\left[3+\left(-1+\rho_G\right)k\left(-3+\rho_G+\left(-1+\rho_G\right)k\right)\right] \tag{38a}$$

$$\kappa_{\Delta Y}=\frac{\left\langle \Delta Y^4\right\rangle}{\left\langle \Delta Y^2\right\rangle^2}=3\frac{(1+k)\left[3-\left(1-\rho_G\right)k\left(-3+\rho_G+\left(-1+\rho_G\right)k\right)\right]}{k\left(1+\left(1-\rho_G\right)k\right)^2}. \tag{38b}$$

Note that according to (38b) $\kappa_{\Delta Y}$ does not depend on θ. Figure 5 depicts the excess kurtosis of ΔY and of Y' versus ρ_G for six values of *k*. At small lags (large ρ_G) $\kappa_{\Delta Y}-3$ exceeds zero by a significant margin, even at large values of *k*. Excess kurtosis decreases as ρ_G decreases (lag increases), rendering the peak of $f_{\Delta Y}$ less sharp and its tails lighter. When $k\geq 10$, the asymptotic value of $\kappa_{\Delta Y}-3$ at large lags is very small (<< 1) and $f_{\Delta Y}$ is virtually Gaussian. When $k>1$, $f_{\Delta Y}$ at small lags has sharper peaks and heavier tails than does $f_{Y'}$, the opposite being true at large lags. Figure 6 depicts on arithmetic and semi logarithmic scales $f_{\Delta Y}$ for $\sigma_G=1.0$, k = 10 and θ = 0.2 and three values of ρ_G. Also shown for comparison is a Gaussian distribution having the same mean and variance as ΔY. As noted in the case of the log-normal subordinator, $f_{\Delta Y}$ exhibits sharp peak and heavy tails when ρ_G = 0.99 (at small lag) and tends to the Gaussian distribution as ρ_G decreases (lag increases).

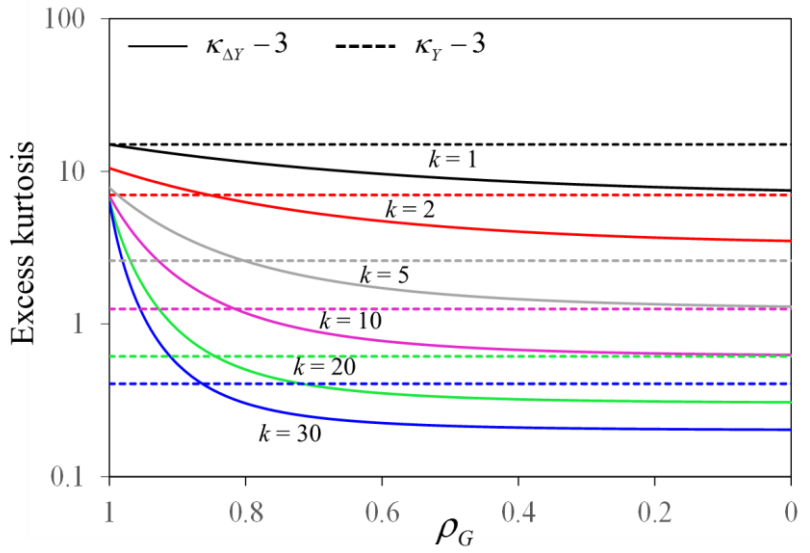

Figure 5. Excess kurtosis of ΔY (continuous curves) and of Y' (horizontal dashed lines) versus ρ_G for six values of *k*.

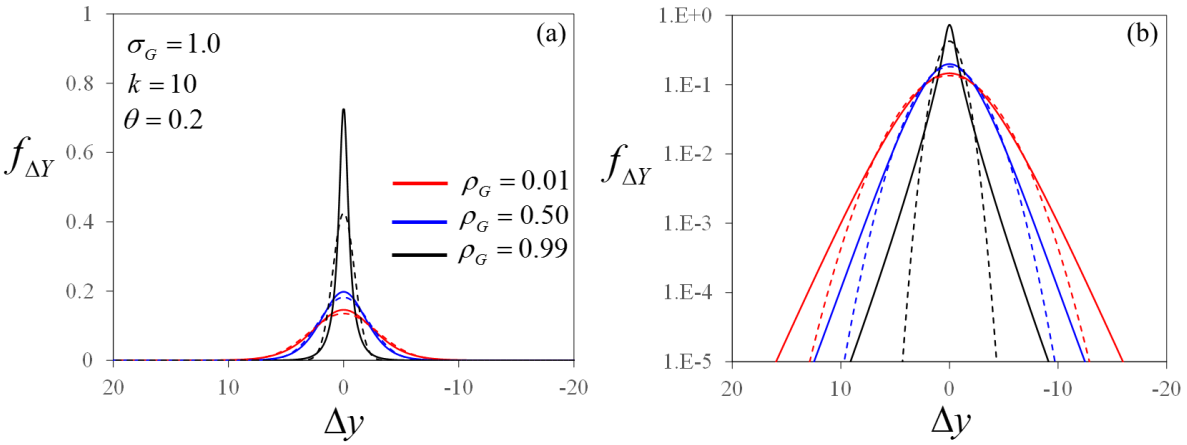

Figure 6. $f_{\Delta Y}$ (36) on (*a*) arithmetic and (*b*) semi logarithmic scales for $\sigma_G=1.0$, k = 10, θ = 0.2 and three values of ρ_G (continuous curves). Also shown are Gaussian distributions having the same mean and variance as ΔY (dashed curves).

The *q*-th order moments of ΔY and $|\Delta Y|$ are given respectively by

$$\left\langle \Delta Y^q \right\rangle = \frac{\sigma_G^q}{\sqrt{2\pi}} 2^{\frac{q-1}{2}} \frac{\left(1+(-1)^q\right)}{\Gamma^2(k)\theta^{2k}} \Gamma\left(\frac{q+1}{2}\right) \int_0^\infty \int_0^\infty \left(u_1^{\,2}+u_2^{\,2}-2\rho_G u_1 u_2\right)^{\frac{q}{2}} u_1^{\,k-1} u_2^{\,k-1} e^{-\frac{1}{\theta}(u_1+u_2)} du_2 du_1 \tag{39}$$

$$\left\langle |\Delta Y|^q \right\rangle = \frac{\sigma_G^q}{\sqrt{2\pi}} 2^{\frac{q+1}{2}} \Gamma\left(\frac{q+1}{2}\right) \frac{1}{\Gamma^2(k)\theta^{2k}} \int_0^\infty \int_0^\infty u_1^{\,k-1} u_2^{\,k-1} \sqrt{\left(u_1^{\,2}+u_2^{\,2}-2\rho_G u_1 u_2\right)^q} e^{-\frac{1}{\theta}(u_1+u_2)} \, du_2 du_1 \tag{40}$$

The variogram of Y' can be obtained directly from (37) as

$$\gamma_Y = \frac{\langle \Delta Y^2 \rangle}{2} = \sigma_G^2 \frac{\Gamma(1+k)}{\Gamma(k)} \theta^2 \left(1+\left(1-\rho_G\right)k\right) = \sigma_G^2 k\theta^2 \left(1+\frac{\gamma_G}{\sigma_G^2}k\right). \quad (41)$$

Note that γ_Y includes a nugget effect (a constant independent of the lag). From (41) and (33) we obtain an expression for the covariance of Y',

$$C_Y(0) = \sigma_G^2 \theta^2 k(1+k), \quad C_Y(s>0) = C_G k^2 \theta^2. \quad (42)$$

This in turn yields the integral scale of Y' to be

$$I_Y = I_G k^2 \theta^2 \quad (43)$$

It is thus seen that a gamma subordinator can either dampen (but not destroy) or amplify the covariance structure of *G*, depending on the set of parameters (*k*, θ) characterizing it. Since *k* and θ can take only strictly positive (non-zero) values, then $I_Y \to I_G$ only for very small values of both of them (i.e., small values of the variance of *U*).

3. Parameter estimation methods

A straightforward way to infer model parameters from a dataset relies on the so-called *Method of moments*, which takes advantage of the analytical expressions of the statistical moments of Y' and ΔY provided in Section 2. The method allows to obtain explicit estimates of the model parameters in replacing the ensemble moments by their sample counterparts. In the following, we detail the methodology developed for the log-normal and for the Gamma subordinator.

3.1 Log-normal subordinator

Parameter Estimation Method a

Method a relies on the marginal frequency distribution and moments of Y', which, for the log-normal subordinator, depend only on two parameters, α and σ_G. One therefore cannot estimate ρ_G by this method. Explicit estimates $\tilde{\alpha}_a$ and $\tilde{\sigma}_{G\_a}$ of parameters α and σ_G can be obtained by replacing the second and fourth moments, $\langle Y'^2 \rangle$ and $\langle Y'^4 \rangle$, of Y' in (20) and (21a) by their sample counterparts, M_2^Y and M_4^Y. This leads to the following expressions:

$$\tilde{\alpha}_a = 2 - \sqrt{\frac{1}{4}\ln\frac{M_4^Y}{3\left(M_2^Y\right)^2}} \tag{44}$$

$$\tilde{\sigma}_{G\_a} = \sqrt{\frac{M_2^Y}{e^{2(2-\tilde{\alpha}_a)^2}}}\,. \tag{45}$$

Parameter Estimation Method b

Method b allows to obtain estimates of all three parameters α, σ_G and ρ_G characterizing Y' and ΔY by relying jointly on samples of both functions. Replacing $\left\langle Y'^2 \right\rangle$, $\left\langle \Delta Y^2 \right\rangle$ and $\left\langle \Delta Y^4 \right\rangle$ in (20), (24) and (25a) by their sample counterparts M_2^Y, $M_2^{\Delta Y}$ and $M_4^{\Delta Y}$ provides explicit estimates $\tilde{\alpha}_b$, $\tilde{\sigma}_{G\_b}$ and $\tilde{\rho}_G$ of the three parameters as

$$\tilde{\alpha}_b = 2 - \sqrt{\ln x} \tag{46}$$

$$\tilde{\sigma}_{G\_b} = \sqrt{\frac{M_2^Y}{e^{2(2-\tilde{\alpha}_b)^2}}} \tag{47}$$

$$\tilde{\rho}_G = e^{(2-\tilde{\alpha}_b)^2} - \frac{M_2^{\Delta Y}}{2\tilde{\sigma}_{G\_b}^2 e^{(2-\tilde{\alpha}_b)^2}}, \tag{48}$$

where

$$d = 2\left[1 - \left(\frac{M_2^{\Delta Y}}{2M_2^Y}\right)^2\right], \quad f = 1 - \frac{M_4^{\Delta Y}}{6\left(M_2^Y\right)^2}, \quad x = \left[\frac{d + \sqrt{d^2 - 4f}}{2}\right]^{\frac{1}{2}}. \tag{49}$$

The joint use of Y' and ΔY is therefore recommended since it allows to (*a*) estimate parameter ρ_G to diagnose the dependency of increment statistics on the lag, and (*b*) enlarge the set of data on which sample moments are computed. With this methodology, it is then possible to obtain one set of $\left(\tilde{\alpha}, \tilde{\sigma}_G, \tilde{\rho}_G\right)$ estimates for each investigated lag. According to our theoretical framework, it is expected that values of $\tilde{\alpha}, \tilde{\sigma}_G$ remain (approximately) constant with lag.

3.2 Gamma subordinator

We start by rewriting (32) as

$$f_{Y'}(y)=\frac{1}{\sqrt{2\pi}\Gamma(k)\left(\theta\sigma_G\right)^k}\int_0^\infty x^{k-2}e^{-\frac{x}{\theta\sigma_G}-\frac{1}{2}\frac{y^2}{x^2}}dx\,, \tag{50}$$

where, $x=u_2\sigma_G$. In a similar way, we rewrite (36) as

$$f_{\Delta Y}(\Delta y)=\frac{1}{\sqrt{2\pi}\Gamma^2(k)\left(\theta\sigma_G\right)^{2k}}\int_0^\infty\int_0^\infty \frac{z^{k-1}x^{k-1}e^{-\frac{1}{\theta\sigma_G}(x+z)-\frac{1}{\left(x^2+z^2-2\rho_G xz\right)}\frac{\Delta y^2}{2}}}{\sqrt{\left(x^2+z^2-2\rho_G xz\right)}}dxdz\,, \tag{51}$$

where $z=u_1\sigma_G$, and $x=u_2\sigma_G$. It is then clear from (50) and (51) that parameters θ and σ_G cannot be independently estimated. In the following we apply the method of moments to estimate the following three parameters: *k*, $C=\theta\sigma_G$ and ρ_G.

Parameter Estimation Method a

Method a relies on the marginal frequency distribution and moments of Y' which depend only on two parameters, *k* and *C*. One therefore cannot estimate ρ_G by this method. Explicit estimates $\tilde{k}_a$ and $\tilde{C}_a$ of parameters *k* and *C* can be obtained by replacing the second and fourth moments, $\langle Y'^2\rangle$ and $\langle Y'^4\rangle$, of Y' in (33) and (34a) by their sample counterparts, M_2^Y and M_4^Y. This leads to the following expressions:

$$\tilde{k}_a=\frac{-\left[15\left(M_Y^2\right)^2-M_Y^4\right]+\sqrt{9\left(M_Y^2\right)^4+\left(M_Y^4\right)^2+42M_Y^4\left(M_Y^2\right)^2}}{2\left[3\left(M_Y^2\right)^2-M_Y^4\right]} \tag{52}$$

$$\tilde{C}_a=\sqrt{\frac{M_Y^2}{k(1+k)}}\,. \tag{53}$$

Parameter Estimation Method b

Method b provides estimates of all three parameters *k*, *C* and ρ_G characterizing Y' and ΔY by relying jointly on sample moments of both functions. Replacing $\langle Y'^2\rangle$, $\langle \Delta Y^2\rangle$ and

$\left\langle \Delta Y^4 \right\rangle$ in (33), (37) and (38a) by their sample counterparts M_2^Y, $M_2^{\Delta Y}$ and $M_4^{\Delta Y}$, provides explicit estimates $\tilde{k}_b$, $\tilde{C}_b$ and $\tilde{\rho}_G$ of the three parameters as

$$\tilde{C}_b = \sqrt{\frac{M_2^Y}{\tilde{k}_b\left(1+\tilde{k}_b\right)}} \tag{54}$$

$$\tilde{\rho}_G = \left(-\frac{M_2^{\Delta Y}}{2M_2^Y}\frac{\left(1+\tilde{k}_b\right)}{\tilde{k}_b} + \frac{1}{\tilde{k}_b} + 1 \right) \tag{55}$$

$$M_4^{\Delta Y} = 12C^4\tilde{k}_b\left(1+\tilde{k}_b\right)\left[3+\left(-1+\tilde{\rho}_G\right)\tilde{k}_b\left(-3+\tilde{\rho}_G+\left(-1+\tilde{\rho}_G\right)\tilde{k}_b\right)\right] \tag{56}$$

Equations (54)-(56) allow to write a third order equation in $\tilde{k}_b$

$$a\,\tilde{k}_b^3 + b\tilde{k}_b^2 + c\tilde{k}_b + d = 0 \tag{57}$$

where

$$\begin{aligned} a &= \frac{M_4^{\Delta Y}}{12\left(M_2^Y\right)^2} - \frac{\left(M_2^{\Delta Y}\right)^2}{4\left(M_2^Y\right)^2}, \qquad b = \frac{M_4^{\Delta Y}}{12\left(M_2^Y\right)^2} - 3\frac{\left(M_2^{\Delta Y}\right)^2}{4\left(M_2^Y\right)^2}, \\ c &= \frac{M_2^{\Delta Y}}{M_2^Y} - 2 - 3\frac{\left(M_2^{\Delta Y}\right)^2}{4\left(M_2^Y\right)^2}, \qquad d = \frac{M_2^{\Delta Y}}{M_2^Y} - 1 - \frac{\left(M_2^{\Delta Y}\right)^2}{4\left(M_2^Y\right)^2}. \end{aligned} \tag{58}$$

Note that only the solution $\tilde{k}_b > 0$ of (57) is acceptable.

4. Application to field data

We conclude this report with a pilot application of the derived models to field neutron porosity data. We let Y represent neutron porosity data from a deep vertical borehole in southwestern Iran recently analyzed by Dashtian et al. (2011), Guadagnini et al. (2015) and Riva et al. (2015). The well is drilled in the Maroon field within which gas drive is used to produce oil and natural gas. A large number (3,567) of neutron porosity data taken at a distance of about 15 cm apart are available, having sample mean M_1^Y = 14% and sample standard deviation 6.4%. Figure 7 plots excess kurtosis of porosity increments, ΔY, versus lag, which ranges from 15 cm to l/2 where l = 543 m is the total depth of the well segment along which data are available. Excess kurtosis $\kappa_{\Delta Y} - 3$ is significantly larger than zero at small lags, then decreases with increasing lags to oscillate around relatively small values (<< 1) at the largest

lags. Included in Figure 7 is a horizontal line denoting excess kurtosis, $\kappa_Y - 3 = 0.64$, of mean-removed porosities $Y' = Y - M_1^Y$. Whereas at small lags $\kappa_{\Delta Y} - 3 > \kappa_Y - 3$, implying that frequency distributions of ΔY exhibit sharper peaks and heavier tails than does that of Y', the opposite happens at large lags. An analogous behavior is predicted by our theoretical models, as highlighted in Figure 1 for a log-normal subordinator with $\alpha > 1$ and in Figure 5 for a gamma subordinator with $k > 1$.

4.1 Log-normal subordinator

By applying *method a* [using (44) and (45)] to this porosity dataset, we find $\tilde{\alpha}_a = 1.78$, $\tilde{\sigma}_{G\_a} = 6.10$ %. Sample distributions of porosity increments are symmetric with peaks and tails that decay with lag. This is illustrated by Figure 8, which displays the pdfs of ΔY evaluated at two lags: $s = 1.22$ m and $s = 7.31$ m. Parameter estimates $\tilde{\alpha}_b$ and $\tilde{\sigma}_{G\_b}$ obtained by *method b*, [using (46)-(48)] plotted respectively versus lag number (1 lag = 0.15 m) in Figures 9a and 9b, oscillate for lags ranging from 1 to 1000 in an irregular fashion about mean values of $\tilde{\alpha}_b$ = 1.75 and $\tilde{\sigma}_{G\_b}$ = 6.15%. These mean values, characterized by small coefficients of variation (0.05 and 0.07, respectively), are very close to their counterparts obtained by *method a*. Estimates of ρ_G also obtained by *method b* are depicted versus lag in Figure 11. Figure 8 also highlights that the analytical pdf computed according to (23) with the set of parameters $\left(\tilde{\alpha}_b, \tilde{\sigma}_{G\_b}, \tilde{\rho}_G\right)$ is in good agreement with sample pdfs for both lags $s = 1.22$ m and $s = 7.31$ m.

4.2 Gamma subordinator

By applying *method a* [using (52) and (53)] we obtained $\tilde{k}_a = 19.1$ and $\tilde{C}_a = 0.33$. Figure 10 depicts the way estimates of *k* and *C* depend on lag number (1 lag = 0.15 m) when they are inferred with *method b* [using (54)-(58)]. We note that *method b* does not yield estimates *k* or *C* which could be considered as constant. Parameter estimates $\tilde{k}_b$ and $\tilde{C}_b$ obtained by *method b*, vary significantly with lag, their mean values being $\tilde{k}_b = 47.6$ and $\tilde{C}_b = 0.52$. The latter, deviate by values estimated by *method a*. We can also note that *k* is almost constant for lags ranging (approximately) between 10 and 80: within such range, mean value and coefficient of variation of $\tilde{k}_b$ are respectively 6.69 and 0.37.

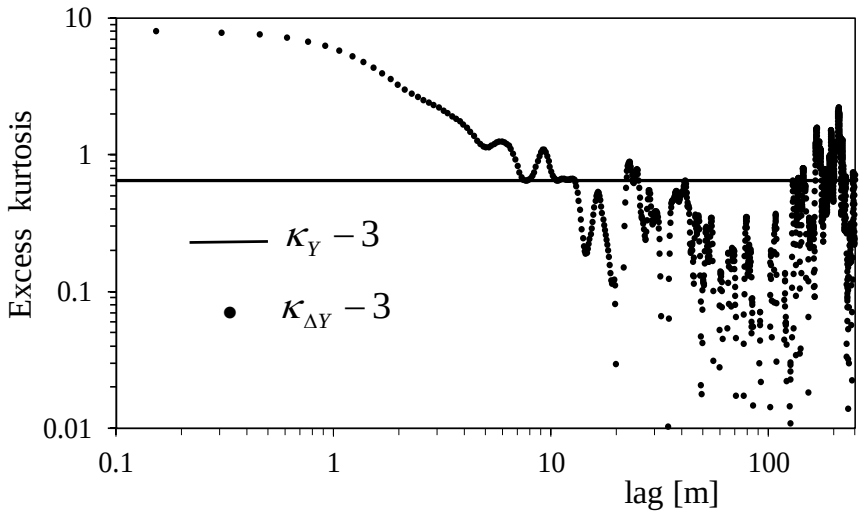

Figure 7. Excess kurtosis of mean-removed porosities data (continuous line) and of porosity increments (symbols) versus lag.

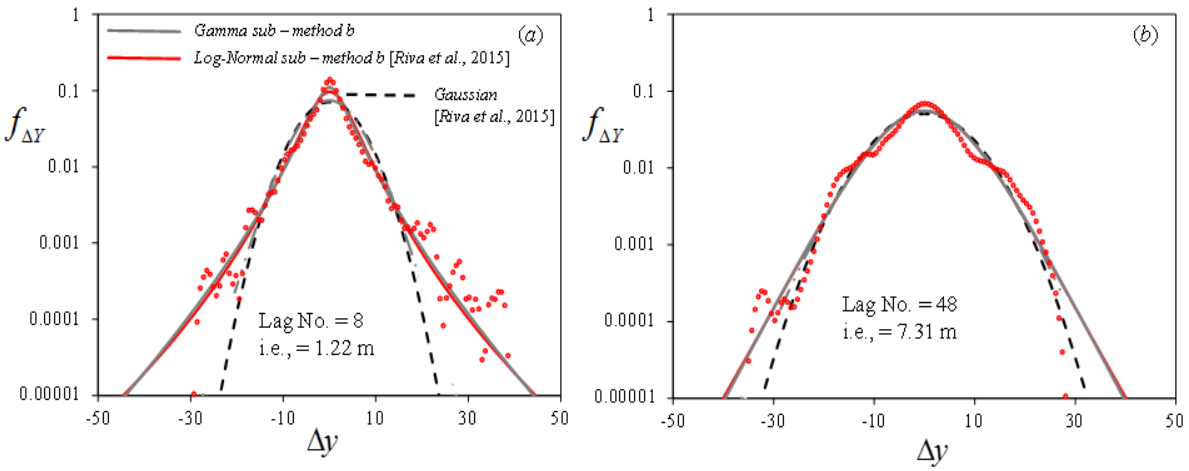

Figure 8. Sample pdf of increments of neutron porosity data, ΔY , (red dots) at two lags: (a) s = 1.22 m and (b) s = 7.31 m. Also shown are Gaussian pdfs with variance equal to that of the sample (dashed curves), $f_{\Delta Y}(\Delta y)$ (23) evaluated using $\tilde{\alpha}_b$, $\tilde{\sigma}_b$, $\tilde{\rho}_G$ (solid red curve) and $f_{\Delta Y}(\Delta y)$ (36) evaluated using $\tilde{k}_b$, $\tilde{C}_b$, $\tilde{\rho}_G$ (solid grey curve)

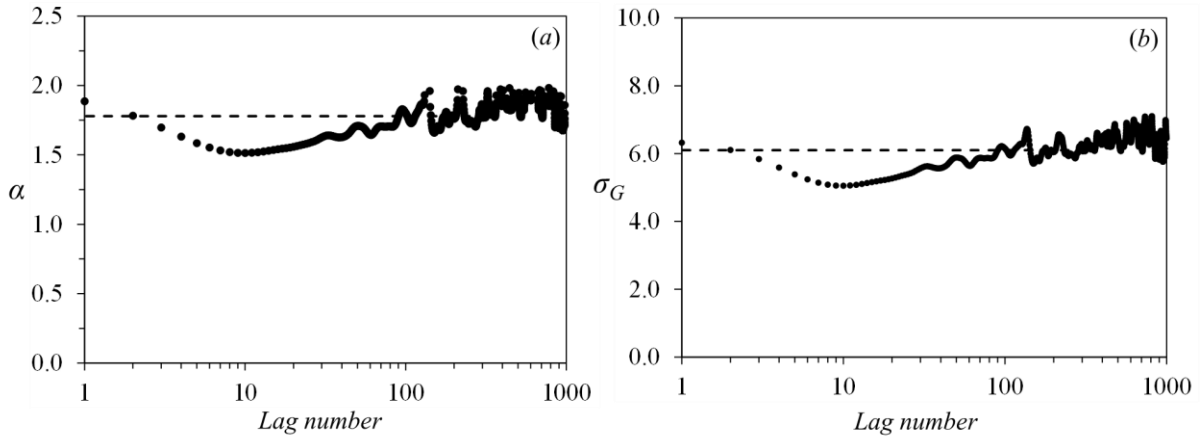

Figure 9. Estimates of the log-normal subordinator parameters (*a*) α and (*b*) σ_G versus lag (1 lag = 0.15 m) computed by (44)-(45) (dashed lines, *method a*) and (46)-(48) (symbols, *method b*).

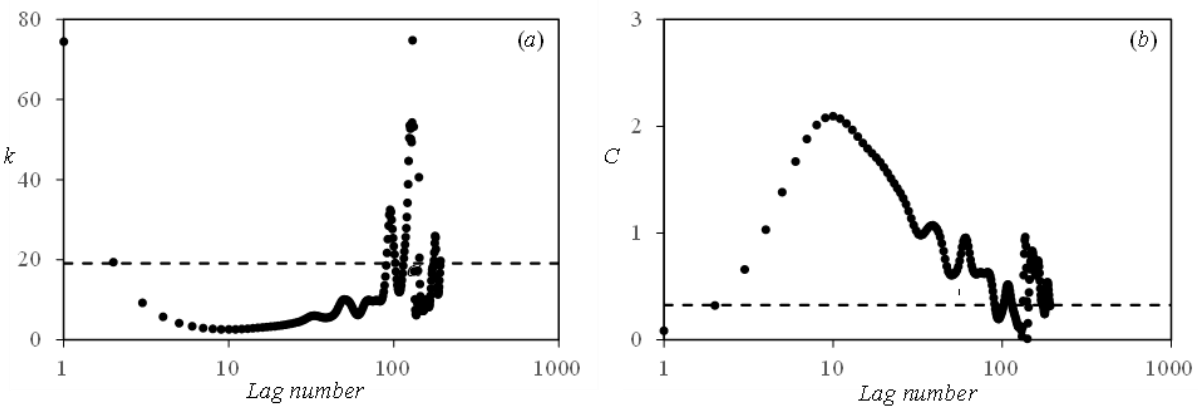

Figure 10. Estimates of the gamma subordinator parameters (*a*) *k* and (*b*) *C* versus lag (1 lag = 0.15 m) computed by (52)-(53) (dashed lines, *method a*) and (54)-(58) (symbols, *method b*).

Figure 8 reports the analytical pdf computed according to (36) with the set of parameters $\left(\tilde{k}_b, \tilde{C}_b, \tilde{\rho}_G\right)$ obtained from (54)-(58). For both lags $s = 1.22$ m and $s = 7.31$ m, (36) is in good agreement with the sample pdf and it is almost overlapped to (23), i.e. to the solution obtained with the log-normal subordinator.

Finally, Figure 11 depicts estimates of ρ_G obtained using a log-normal and a gamma subordinator versus lag. It can be noticed that the shape of the correlation function of the Gaussian process is quite insensitive to the choice of subordinator we employ. This might imply that the signature of the correlation imprinted by the Gaussian process can be considered as a distinct feature of the system.

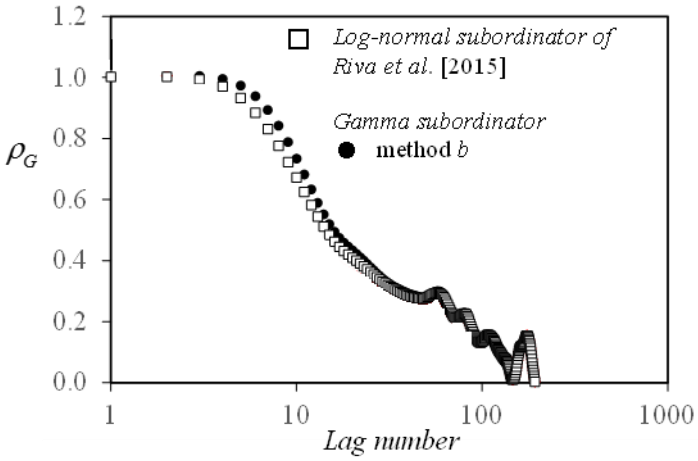

Figure 11. Estimates of ρ_G versus lag obtained using a log-normal and a gamma subordinator.

References

Dashtian, H., G. R. Jafari, M. Sahimi, and M. Masihi (2011), Scaling, multifractality, and long-range correlations in well log data of large-scale porous media, Physica A, 390, 2096-2111, doi:10.1016/j.physa.2011.01.010.

Clark, P.K. (1973), A subordinated stochastic process model with finite variance for speculative prices, *Econometrica*, 41, 1.

Guadagnini, A., S.P. Neuman, M.G. Schaap, and M. Riva (2013) Anisotropic statistical scaling of vadose zone hydraulic property estimates near Maricopa, Arizona, *Water Resour. Res*., 49, doi:10.1002/2013WR014286.

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

Guadagnini, A., S. P. Neuman, T. Nan, M. Riva, and C. L. Winter (2015), Scalable statistics of correlated random variables and extremes applied to deep borehole porosities, *Hydrology and Earth System Sciences*, Hydrol. Earth Syst. Sci., 19, 1–17, doi:10.5194/hess-19-1-2015

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

Meerschaert, M.M., T.J. Kozubowski, F.J. Molz, and S. Lu (2004), Fractional Laplace model for hydraulic conductivity. *Geophys Res Lett* 31, L08501.

Painter, S. (1996), Evidence for non-Gaussian scaling behavior of heterogeneous sedimentary formations, *Water Resour. Res*., 32(5), 1183-1195.

Painter, S. (2001), Flexible scaling model for use in random field simulation of hydraulic conductivity, *Water Resour. Res*., 37, 1155-1163.

Panzeri, M., M. Riva, A. Guadagnini, and S.P. Neuman (2016), Theory and generation of conditional, scalable sub-Gaussian random fields, *Water Resour. Res*., 52, doi: 10.1002/2015WR018348.

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

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

Riva, M., S.P Neuman and A. Guadagnini (2015). New scaling model for variables and increments with heavy-tailed distributions. Water Resour. Res. 51(6), 4623-4634.

Siena, M., A. Guadagnini, M. Riva, and S. P. Neuman (2012), Extended power-law scaling of air permeabilities measured on a block of tuff, *Hydrol. Earth Syst. Sci*., 16, 29-42, doi:10.5194/hess-16-29-2012.

Siena, M., A. Guadagnini, M. Riva, B. Bijeljic, J. P. Pereira Nunes, and M. J. Blunt (2014), Statistical scaling of pore-scale Lagrangian velocities in natural porous media, *Phys. Rev. E*, 90, 023013, doi: 10.1103/PhysRevE.90.023013.