Application of Lévy Processes in Modelling (Geodetic) Time Series With Mixed Spectra

Jean-Philippe Montillet [1,2], Xiaoxing He [3], Kegen Yu [4], and Changliang Xiong [5]

[1] Physikalisch - Meteorologisches Observatorium Davos / World Radiation Center, Davos, Switzerland
[2]Space and Earth Geodetic Analysis laboratory (SEGAL), University Beira Interior, Covhila, Portugal
[3]School of Civil Engineering and Architecture, East China Jiaotong University, Nan Chang, China
[4]School of Environmental Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, China
[5]Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Science (CAS), Xiaohongshan West Road, Wuhan, China

Correspondence: J.-P. Montillet (jpmontillet@segal.ubi.pt)

Abstract. Recently, various models have been developed, including the fractional Brownian motion (fBm), to analyse the stochastic properties of geodetic time series, together with the estimated geophysical signals. The noise spectrum of these time series is generally modelled as a mixed spectrum, with a sum of white and coloured noise. Here, we are interested in modelling the residual time series, after deterministically subtracting geophysical signals from the observations. This residual time series is then assumed to be a sum of three stochastic processes, including the family of Lévy processes. The introduction of a third stochastic term models the remaining residual signals and other correlated processes. Via simulations and real time series, we identify three classes of Lévy processes: Gaussian, fractional and stable. In the first case, residuals are predominantly constituted of short-memory processes. Fractional Lévy process can be an alternative model to the fBm in the presence of long-term correlations and self-similarity property. Stable process is here restrained to the special case of infinite variance, which can be only satisfied in the case of heavy-tailed distributions in the application to geodetic time series. Therefore, it implies potential anxiety in the functional model selection where missing geophysical information can generate such residual time series.

1 Introduction

Among the geodetic data, Global Navigation Satellite System (GNSS) time series have been of particular interest for the study of geophysical phenomenon at regional and global scales (e.g., study of the crustal deformation due to large Earthquakes, sea-level rise (Bock and Melgare , 2016; Herring et al. , 2016; He et al. , 2017)). This time series are the estimated daily position of the receiver coordinates. The position vector of a station can be decomposed in a geocentric cartesian axis system or in a local or topocentric cartesian axis system (E,N,U) in which the axes point east, north and up. These coordinates are influenced by the sum of three displacement modes (distinct classes of motion) that describes the progressive ground motion, any instantaneous jumps in position, and periodic or cyclical displacements. The progressive ground motion is generally refers as the tectonic rate. Jumps include coseismic offsets, which are real movements of the ground, and artificial offsets associated

with changes in the GNSS antenna and/or its radome, or changes in the antenna monument, etc. Nearly all GNSS time series exhibit a seasonal cycle of displacement which can be modelled as a Fourier series. These cycles are caused by seasonal changes in the water, snow and ice loads supported by the solid earth or (less commonly) by seasonal changes in atmospheric pressure. Therefore the model associated with each class of motion (or geophysical signals) is here defined as a functional model following Bevis and Brown (2014) and Montillet and Bos (2019) [Chapter 1]. Furthermore, these time series contain white noise and coloured noise. To model the different noise components, a stochastic noise model is defined. To name a few, it includes the First Order Gauss-Markov (FOGM) model, the white noise with power-law noise (Williams , 2003; Williams et al. , 2004), the Generalized Gauss Markov noise model (GGM), or the Band-pass noise (Langbein , 2008; Langbein and Svarc , 2019). (e.g., Flicker noise and white noise). The scientific community agrees with the existence of a trade-off in estimating both the stochastic and functional models (He et al. , 2017). More precisely, the choice of the stochastic model directly influences the estimation of the geophysical signals included in the functional model (i.e., tectonic rate, seasonal variations, slow-slip events (Bock and Melgare , 2016; He et al. , 2017)).

In addition, recent studies (Langbein and Svarc , 2019; He et al. , 2019) have also advocated the introduction of a random-walk to model small jumps and residual transient signals which is a non-stationary stochastic process. Thus, several studies (Montillet and Yu , 2015) proposed the use of the fractional Brownian motion (fBm), first developed by Mandelbrot et al. (1968), in order to model long-memory processes. Botai et al. (2011) and Montillet and Yu (2015) focused on modelling (residual) geodetic time series using the family of Lévy α-stable distributions (Samorodnitsky and Taqqu , 1994; Nolan , 2018). The application of this family of distribution was supported by the ability to model long-memory processes and the existence of impulsive signals/noise bursts in the data sets suggesting deviations from Gaussian distribution (Botai et al. , 2011).

This work discusses several statistical assumptions (i.e. stationary properties, presence of long-term correlations) on the underlying processes in the GNSS time series, justifying the application of the fBm and the family of Lévy α-stable distributions introduced by Montillet and Yu (2015). The Lévy stable distributions can model the heavy tail characteristics of some data sets with generally infinite variance. For example, the presence of unmodelled large jumps within the data can produce a distribution with large tails and infinite variance. In order to take into account a large variety of scenarios, we investigate and identify within the family of Lévy processes, which process can be applied to model geodetic time series.

Here, the statistical modelling is applied to residual time series following Montillet and Yu (2015). The residual time series are defined as the remaining time series after subtracting deterministically modelled tectonic rate and seasonal components (i.e. the functional model), from the GNSS observations. Therefore, our assumption is that the family of Lévy processes can model the remaining geophysical signals and correlations which have not been captured by the initial model used to produce the residual time series.

The next section starts with the statistical inference on the residual geodetic time series, including the application of the fBm model and the relationship with the Fractional Autoregressive Integrated Moving Average (FARIMA) model. Section 2.3 presents the assumptions on the use of the Lévy processes in the model of the residual time series. To do so, we model the residual geodetic time series as a sum of three stochastic processes, with the hypothesis that the third one is a Lévy process. It involves some justifications compared with established models in the scientific community. In Section 3, we develop an

N-step method based on the variations of the stochastic and functional models when varying the time series' length. This algorithm should verify our statistical assumptions on the third process. Section 3.1.1 and Section 3.1.2 focus on the application to simulated and real time series. Finally, Section 3.2 discusses the limits of modelling geodetic time series with Lévy processes.

2 The Stochastic Properties of the Residual Time Series and Statistical Inferences

2.1 Stochastic Modelling of Residual GNSS Time Series

Let us model the GNSS observations and residual time series as an additive model:

$$\begin{aligned} x_0(t_i) &= s_0(t_i) + n(t_i) \\ x(t_i) &= s_r(t_i) + n(t_i) \\ s_r(t_i) &= x_0(t_i) - s_0(t_i) \end{aligned} \tag{1}$$

x_0 is the time series defined as the GNSS observations, x the residual time series after subtracting the functional model (s_0). At each i-th epoch, $x(t_i)$ is a sum of a residual geophysical signal $s_r(t_i)$ and noise $n(t_i)$. Following Williams (2003) and He et al. (2017), the spectrum of the (residual) GNSS time series is best characterised by a stochastic process following a power-law with index K (i.e. $P(f) = P_0(f/f_s)^K$, f is the frequency, P_0 is a constant, f_s the sampling frequency). A power-law noise model means that the frequency spectrum is not flat but is governed by long-range dependencies. An example is shown in Figure 3 using the $ASCO$ station, other examples are displayed in supplementary material. Power-law noise is a type of coloured noise. The coloured noise results from various parameters during the processing of the GNSS observations such as the mismodelling of GNSS satellites orbits, Earth orientation parameters, large-scale atmospheric or hydrospheric effects (Williams , 2003; Klos et al. , 2018). The stochastic noise model of the (residual) GNSS time series is then described with the variance:

$$E\{\mathbf{n}^T\mathbf{n}\} = \sigma_{wn}^2\mathbf{I} + \sigma_{pl}^2\mathbf{J}(K) \tag{2}$$

where the vector $\mathbf{n} = [n(t_1), n(t_2), ..., n(t_L)]$ is a multivariate noise and continuous process. At each epoch, we define $n(t_i) = n_{wn}(t_i) + n_{pl}(t_i)$, with $n_{wn}(t_i)$ and $n_{pl}(t_i)$ the white Gaussian noise (zero mean) and the power-law noise sample respectively. Note that this type of time series belongs to the family of mixed spectra, where the mixed spectrum results from the sum of the spectra corresponding to the two kinds of noise (Li , 2013). T is the transpose operator, $\mathbf{I}$ the identity matrix, σ_{pl}^2 the variance of the power-law noise and $\mathbf{J}(K)$ the covariance matrix of the power-law noise (K in $]0, 2]$). The definition of $\mathbf{J}$ depends on the assumptions on the type of coloured noise (see supplementary material).

We estimate jointly the functional and stochastic models in order to produce x, based on a maximum likelihood estimator (MLE). To recall Montillet and Bos (2019) (Chapter 2), for linear models, the log-likelihood for a time series of length L can be rewritten as:

$$\ln(Lo) = -\frac{1}{2}\left[L\ln(2\pi) + \ln(\det(\boldsymbol{C})) + (\boldsymbol{x}_0 - \boldsymbol{A}\boldsymbol{z})^T\boldsymbol{C}^{-1}(\boldsymbol{x}_0 - \boldsymbol{A}\boldsymbol{z})\right] \tag{3}$$

This function must be maximised. Assuming that the covariance matrix $\boldsymbol{C}$ is known, then it is a constant and does not influence finding the maximum. $\boldsymbol{C}$ is here defined by Eq. (2). The term $(\boldsymbol{x}_0 - \boldsymbol{Az})$ represent the observations minus the fitted model or $\boldsymbol{x}$ in Eq. (1). Note that $(\boldsymbol{Az})$ is the matrix notation of s_0. The last term can be written as $\boldsymbol{x}^T\boldsymbol{C}^{-1}\boldsymbol{x}$ and it is a quadratic function, weighted by the inverse of matrix $\boldsymbol{C}$. To select the geophysical model, and therefore estimate the associated parameters, one needs to consider carefully the location of the GNSS stations and the surrounding geodynamics. The model of s_0 is discussed in the supplementary material together with the software used to carry out the maximisation of $\ln(Lo)$. Note that the value of L is here at least 9 years (3285 observations) in order to be able to model correctly the long-range dependencies associated with the coloured noise and to detect slow transient signals according to He et al. (2019).

In the modelling of GNSS time series, a strong assumption is the so-called Gauss-Markov hypothesis (e.g., Montillet and Bos (2019) - *Chapter* 2) which states that the noise is Gaussian distributed. In order to keep applying the Gauss-Markov assumption on the noise observations of geodetic time series, we assume that the mean of the coloured noise is equal to $\mu_C(t)$, slowly varying with time. We then rule out the occurrence of specific events of large amplitude such as aggregations or burst of spikes (i.e. intermittency) which could invalidate such assumption (see Supplementary material for more information).

Moreover, if the probability density function of the noise is Gaussian or has a different density function with a finite value of variance, its fractal properties can be described by the Hurst parameter (H). The authors in Montillet et al. (2013) use the fBm in order to model the statistical properties of the residual time series. The essential features of this process are its self-similar behaviour, meaning that magnified and re-scaled versions of the process appear statistically identical to the original, together with its non-stationary property implying a never-ending growth of variance with time (Mandelbrot et al. , 1968). Previous studies (e.g., Mandelbrot et al. , 1968; Eke et al. , 2002) showed that H is directly connected with K by the relation:

$$K = 2H - 1 \tag{4}$$

With this definition, flicker noise corresponds to K equal to 1 or H equal to 1, random walk to K equal to 2 or H equal to 1.5, and white noise to K equal to 0 (H equal to 0.5). Note that Eq. (4) is established for the fractional Gaussian noise according to (Eke et al. , 2002). The random-walk and the flicker noise are then classified as long-term dependency phenomena (Montillet et al. , 2013).

Long-memory processes are modelled with a specific class of ARIMA models called FARIMA by allowing for non-integer differentiating. A comprehensive literature on the application of FARIMA can be found in financial analysis (Granger and Joyeux , 1980; Panas , 2001) and in geodesy (Li et al. , 2000; Montillet and Yu , 2015; Montillet and Bos , 2019). This model can generate long-memory processes based on the different values of the fractional index d (Granger and Joyeux , 1980). When d equal to 0 it is an Autoregressive Moving Average (ARMA) process exhibiting short memory; when $-0.5 \leq d < 0$ the FARIMA process is said to exhibit intermediate memory or anti-persistence (Pipiras and Taqqu , 2017). This is very similar to the description of H in the fBm. In the supplementary material, we recall the relationship between FARIMA, ARMA and fBm.

2.2 α Stable Random Variable and the Lévy α-Stable Distribution

The fBm and the fractional Lévy distribution are well-known in statistics (Samorodnitsky and Taqqu , 1994) and in financial analysis (Panas , 2001; Wooldridge , 2010). The fractional Lévy distribution can model the Lévy processes and in particular the general family of α stable Lévy processes which can be self-similar and stationary (Samorodnitsky and Taqqu , 1994). Let us recall the definition of a stable random variable.

Definition *(Nolan , 2018), chap. 1, definition, 1.6* A random variable X is stable if and only if $X \stackrel{d}{=} aZ+b$, where $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $a \neq 0$, $b \in \mathbb{R}$ and Z is a random variable with characteristic function $\phi(u) = E\{\exp(iuZ)\} = \int_{-\infty}^{\infty} \exp(iuz)F(z)\,dz$. $F(z)$ is the distribution function of Z. $E\{.\}$ is the expectation operator. The characteristic function is:

$$\phi(u) = \begin{cases} \exp(-|u|^{\alpha}[1 - i\beta \tan\frac{\pi\alpha}{2}(\text{sign}(u))]), & \text{if } \alpha \neq 1 \\ \exp(-|u|[1 + i\beta\frac{2}{\pi}\text{sign}(u)]), & \text{if } \alpha = 1 \end{cases} \tag{5}$$

Where sign is the signum function, α is the characteristic exponent which measures the thickness of the tails of these distributions (the smaller the values of α, the thicker the tails of distribution are), $\beta \in [-1, 1]$ is the symmetry parameter which set the skewness of the distribution. In general, no closed-form expression exists for these distributions, except for the Gaussian (α equal to 2), Pearson (α equal to 0.5, β equal to -1) and Cauchy (α equal to 1, β equal to 0) distributions.

Now, restricting to our case study, we assume that if the stochastic process exhibits a self-similar property, then it can be modeled by the fBm. Following (Weron et al. , 2005), the most commonly used extension of the fBm to the α-stable case is the fractional Lévy stable motion (fLsm). The fLsm is H-self-similar and has stationary increments, with H the Hurst parameter described before. Both the fBm and the fLsm follow an integral representation, with different properties of their kernel (see the supplementary material). The relationship between the fLsm reduces to the fBm when $\alpha = 2$. If $H = 1/\alpha$, we obtain the Lévy α-stable motion which is an extension of the Brownian motion to the α-stable case. Note that the Lévy α-stable motion belongs to the Lévy processes.

2.3 The Residual Time Series Modelled as a Sum of Three Stochastic Processes

The residual time series is now modelled as a sum of three stochastic processes. Namely, it is the sum of a white noise, a coloured noise and a third process. It is a similar approach used in previous works (Langbein , 2008; Davis et al. , 2012; Langbein and Svarc , 2019; He et al. , 2019) looking at the presence of a random-walk component in the stochastic model, hence adding a third covariance matrix in Eq. (2). We postulate that this unknown stochastic process belongs to the Lévy processes, classified in three types depending on the assumptions on the underlying process:

1. (Lévy Gaussian) The Lévy process is a Gaussian Lévy process if the process follows the properties of a pure Brownian motion also called a Wiener process (identity variance matrix, zero-mean, stationary independent increment - (Haykin , 2002; Wooldridge , 2010)). That is the special case of the fLsm and fBm with $H = 1/2$. The residual time series is assumed to contain mostly short-term correlations modelled with an ARMA process. The residual time series should be modelled with a multivariate Gaussian distribution.

2. (Fractional Lévy) The residual time series exhibits self-similarity with possibly long-term correlations. The Fractional Lévy process is described by the model of the fLsm for the specific case reduced to the fBm. The long-term correlation process is mostly due to the presence of coloured noise (He et al. , 2017). As explained in Montillet and Yu (2015), the ratio of the amplitude of the coloured over white noise determines which stochastic model of the residual time series should be the most suitable between the FARIMA and ARMA processes. However, the Gauss-Markov assumption is still valid, therefore the residual time series should be modelled with a multivariate Gaussian distribution.

3. (Stable Lévy) The Lévy process is a Lévy α-stable motion (not reduced to the fBm case). The Gauss-Markov assumption is not holding anymore. The distribution of the residual time series is potentially skewed, not symmetric, with possibly heavy tails, hence modelling with a Lévy α-stable distribution. With the relationship between the Lévy α-stable motion, the fBm and the FARIMA, we assume that the stochastic properties of the residual time series should be described with the FARIMA, especially in the presence of large amplitude coloured noise.

In the application to geodetic time series, the third case occurs mainly due to a misfit between the selected (stochastic and functional) model and the observations. Therefore, the residual time series withholds some remaining unmodelled geophysical signals or unfiltered large outliers which can potentially undermine the Gauss-Markov assumption (e.g., presence of heavy tails in the distribution of the residual time series). For example, if small jumps (or Markov jumps), outliers or other unknown processes are present, it results in a distribution of the residual time series not symmetric and with heavy tails. The functional model describing those jumps is a Heaviside step function (Herring et al. , 2016; He et al. , 2017) as shown in the appendices. In order to assume a Lévy α-stable motion as the underlying stochastic model in geodetic time series, we restrict our assumption to small undetectable offsets, modelling them potentially as random-walk.

2.4 The N-Step Method

To recall Section 2.1, let us describe the functional model and the stochastic noise model described in Eq. (2) as a functional interpretation called $\mathcal{F}(\theta_1)$ and $\mathcal{G}(\theta_2)$. The functional model is the modeled geophysical signals, whereas the stochastic noise model described using the covariance matrix in Eq. (2) is equal to $\mathcal{G}(\theta_2)$. We define $\theta_1 = [a, b, (c_j, d_j)_{j=\{1,N\}}]$ and $\theta_2 = [a_{wh}, b_{pl}, K]$, the vector parameters for the functional and stochastic noise model respectively. For simplification, we have not included in the functional model the estimation of possible offsets in the time series (see appendices for the discussion). Also, a_{wh} and b_{pl} are the amplitude of the white and power-law noise respectively.

Furthermore, our method is based on varying the length of the time series resulting in the variations of the stochastic and functional models, which should allow classifying the type of Lévy process. The variations of the length of the time series should take into account that the coloured noise is a non-stationary signal (around the mean), and thus the properties (i.e. b_{pl}, K) vary non-linearly. However, varying the length of the time series over several years is not realistic taking into account that real time series can record undetectable transient signals, undocumented offsets and other non-deterministic signals unlikely to be modelled precisely (Montillet et al. , 2015). That is why we restrain the variations of the time series length to 1 year.

Let us call the geodetic time series $\mathbf{s}_1 = [s(t_1), ..., s(t_L)]$, $\mathbf{s}_2 = [s(t_1), ..., s(t_{L+1})]$ and $\mathbf{s}_N = [s(t_1), ..., s(t_{L+N})]$ at the first, second and N-th variation respectively. Note that the N samples are equal to 1 year in this example, and for simplification we add only 1 sample at each step. That is not realistic, but the sole purpose is to be a pedagogical example. According to the functional notation above, the GNSS observations $\mathbf{s}$ and the estimated stochastic noise and functional models $\hat{\mathbf{s}}$ are equal to:

$$\begin{aligned} \mathbf{s} &= \mathcal{F}(\theta_1) + \mathcal{G}(\theta_2) \\ \hat{\mathbf{s}} &= \mathcal{F}(\hat{\theta_1}) + \mathcal{G}(\hat{\theta_2}) \end{aligned} \tag{6}$$

Let us describe the method for the first, second and N-th step such as:

$\underline{1^{st}\ step:}$

$$\begin{aligned} \mathbf{s}_1 &= [s(t_1), ..., s(t_L)] \text{ (Time Series)} \\ [\hat{\theta_1}]_1, [\hat{\theta_2}]_1 &= \underset{\theta_1, \theta_2}{argmax} \{\mathbf{s}_1 - (\mathcal{F}(\theta_1) + \mathcal{G}(\theta_2))\} \\ \hat{\mathbf{s}}_1 &= \mathcal{F}([\hat{\theta_1}]_1) + \mathcal{G}([\hat{\theta_2}]_1) \text{ (Estimated model)} \end{aligned}$$

(8)

$$\begin{aligned} \Delta\mathbf{s}_1 &= \mathbf{s}_1 - \mathcal{F}([\hat{\theta_1}]_1) \text{ (residual T.S.)} \\ &\simeq \mathcal{G}([\hat{\theta_2}]_1) + \epsilon_1 \end{aligned}$$

$\underline{2^{nd}\ step:}$

$$\begin{aligned} \mathbf{s}_2 &= [s(t_1), ..., s(t_{L+1})] \text{ (Time Series)} \\ [\hat{\theta_1}]_2, [\hat{\theta_2}]_2 &= \underset{\theta_1, \theta_2}{argmax} \{\mathbf{s}_2 - (\mathcal{F}(\theta_1) + \mathcal{G}(\theta_2))\} \\ \hat{\mathbf{s}}_2 &= \mathcal{F}([\hat{\theta_1}]_2) + \mathcal{G}([\hat{\theta_2}]_2) \text{ (Estimated model)} \end{aligned}$$

(10)

$$\begin{aligned} \Delta\mathbf{s}_2 &= \mathbf{s}_2 - \mathcal{F}([\hat{\theta_1}]_2) \text{ (residual T.S.)} \\ &\simeq \mathcal{G}([\hat{\theta_2}]_2) + \epsilon_2 \end{aligned}$$

$\underline{N-th\ step:}$

$$\begin{aligned} \mathbf{s}_N &= [s(t_1), ..., s(t_{L+N})] \text{ (Time Series)} \\ [\hat{\theta_1}]_N, [\hat{\theta_2}]_N &= \underset{\theta_1, \theta_2}{argmax} \{\mathbf{s}_N - (\mathcal{F}(\theta_1) + \mathcal{G}(\theta_2))\} \\ \hat{\mathbf{s}}_N &= \mathcal{F}([\hat{\theta_1}]_N) + \mathcal{G}([\hat{\theta_2}]_N) \text{ (Estimated model)} \end{aligned}$$

(12)

$$\Delta \mathbf{s}_N = \mathbf{s}_N - \mathcal{F}([\hat{\theta_1}]_N) \text{ (residual T.S.)}$$
$$\simeq \mathcal{G}([\hat{\theta_2}]_N) + \epsilon_N \quad (13)$$

where $\hat{.}$ corresponds to the estimated vector or observations. $[.]_j$ means the j-th iteration of the estimated quantity. $\Delta \mathbf{s}_j$ is the residual time series at the j-th step. ϵ_j (with j in $[1, 2, .., N]$) is the unmodelled signals and stochastic processes at the j-th step. Note that the methodology requires the estimation of the parameters of the functional and stochastic noise models $[\hat{\theta_1}]_j, [\hat{\theta_2}]_j$ via MLE as described in Section 2.1 in the maximization of $\ln(Lo)$ in Eq. (3) (see also Hector software Bos et al. (2013) in the supplementary material).

To recall the assumptions in Section 2.3, the residual time series $\Delta \mathbf{s}_N$ is modelled as a sum of three stochastic processes corresponding to the white noise, coloured noise and a Lévy process. Using N iterations/steps and our statistical inferences on the Lévy processes (i.e., Lévy Gaussian, Fractional Lévy and Stable Lévy), we make several assumptions on the estimated parameters and selected stochastic models in order to characterize the third process. Table 1 summarises these assumptions. We use specific mathematical symbols to differentiate between them. $\triangleq$ means the equality in terms of distribution. $\simeq$, $\sim$ and $\neq$ are related to the variations of the estimated parameters of the stochastic model associated with the first and the N-th step. The symbol $\simeq$ means that there are little differences (less than 3%) between the estimated parameters of the stochastic noise model associated with the first and the N-th iteration. The symbol $\sim$ means that we allow bigger differences up to 20% . With much larger values, we use the symbol $\neq$. Note that the variation of the estimated stochastic noise parameters $[\hat{\theta_2}]_j$ between the first and the j-th step is calculated using the sum of the difference in absolute value between the estimates (e.g., between the first and $j+1$ step, $||[\hat{\theta_2}]_1 - [\hat{\theta_2}]_{j+1}||$). We deduce a percentage of the variations based on the sum in absolute value of the estimates $[\hat{\theta_2}]_1$.

Table 1. Assumptions on the functional model and the stochastic parameters estimated via N iterations (see,N-Step method) to characterize the type of Lévy processes within the geodetic time series. The symbols and notations are explained in Section 2.4.

Type of Process	Lévy Gaussian	Fractional Lévy	Stable Lévy
Mathematical Assumptions	$\mathcal{G}([\hat{\theta_2}]_1) \simeq \mathcal{G}([\hat{\theta_2}]_N)$ $\mathcal{F}([\hat{\theta_1}]_1) \simeq \mathcal{F}([\hat{\theta_1}]_N)$	$\mathcal{G}([\hat{\theta_2}]_1) \sim \mathcal{G}([\hat{\theta_2}]_N)$ $\mathcal{F}([\hat{\theta_1}]_1) \sim \mathcal{F}([\hat{\theta_1}]_N)$	$\mathcal{G}([\hat{\theta_2}]_1) \neq \mathcal{G}([\hat{\theta_2}]_N)$ $\mathcal{F}([\hat{\theta_1}]_1) \neq \mathcal{F}([\hat{\theta_1}]_N)$
(Distribution) $\Delta^1 \mathbf{s} \triangleq$	Gaussian	Gaussian	Lévy α-stable
Model To Characterize Processes	ARMA(p,q)	ARMA(p,q) or FARIMA(p,d,q)	FARIMA(p,d,q)

Furthermore, the fitting of the ARMA(p,q) and FARIMA(p,d,q) model to the residual time series is carried out by maximum-likelihood following Sowell (1991). The lags p and q vary within the interval $[0, 5]$. Also, the selection of the model which

best fits the residual time series, is performed by minimizing the Bayesian Information Criterion (BIC) following Montillet and Yu (2015). Finally, one can wonder if the anxiety in the model selection ARMA, FARIMA) in presence of heavy-tails can modify the performance of the BIC. This topic is currently debated in the statistical community (e.g., (Panahi , 2016)). Large tails should be detected in the fitting of the Lévy α-stable distribution. Various methods exist to estimate the parameters of this distribution (Koutrouvelis , 1980), however we use the maximum-likelihood method of Nikias and Shao (1995). Due to the direct relationship between the index α and H recalled in Section 2.1, we assume that the amplitude of the coloured noise is higher than the white noise, therefore the FARIMA should be chosen de facto over the ARMA model.

3 Lévy Processes Applied to Geodetic Time Series Analysis

This section deals with the application of the N-step algorithm to simulated and real time series. This approach should verify our statistical inferences formulated in Section 2.3. Note that the simulations of the GNSS time series are comprehensively explained in the supplementary material (supplement D).

3.1 Application to Simulated and Real Time Series

We have restrained our simulations to the stochastic model corresponding to the flicker noise (with white noise - $FN+WN$) and power-law (with white noise - $PL+WN$). In addition to simplify our study, we have preliminary applied the method based on the Akaike information criterion developed in He et al. (2019) on the real time series to select the optimal stochastic noise model. Therefore we have selected real time series with stochastic models $FN+WN$ and $PL+WN$. We are not going to develop further this topic in this study, but readers can refer to He et al. (2019).

3.1.1 Simulated Time Series

We simulate 10 years long time series fixing a_{wh} to 1.6 mm; the tectonic parameters a varying between $[1-3]$ mm/yr and b equal 0; and the seasonal signal with only the first harmonic (c_1, e_1) equal to $(0.4, 0.2)$ mm/yr. Details on the noise simulations are given in the supplementary material. According to Table 1, we vary the amplitude of coloured noise b_{cl} following three scenarios:

A. low value (i.e. $b_{cl} < 0.1$ mm/$yr^{K/4}$)

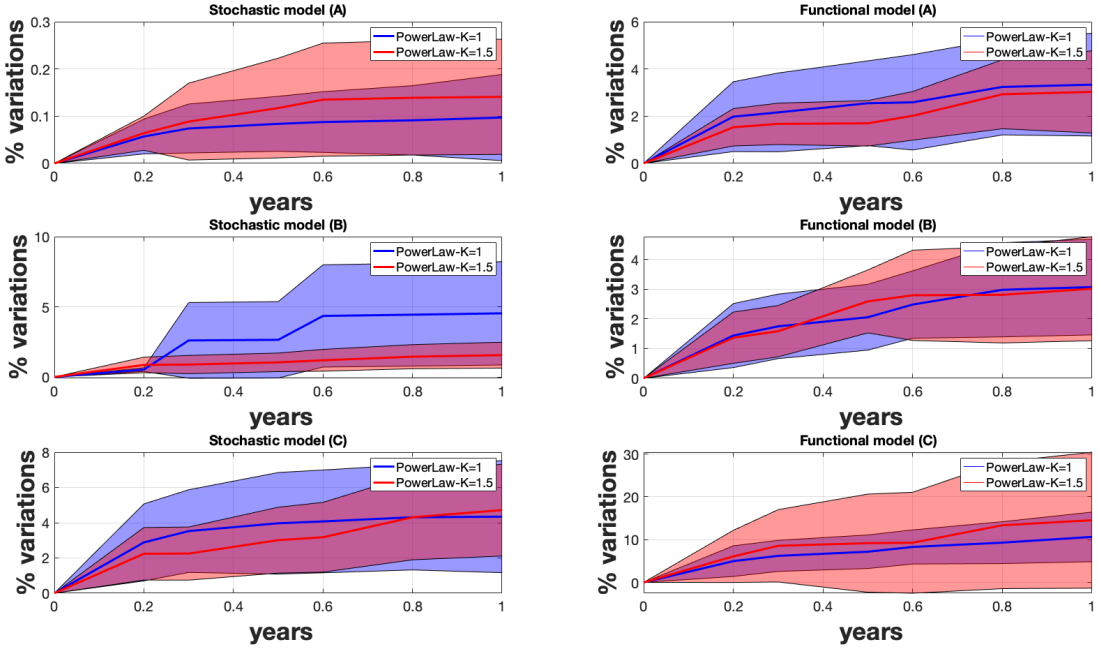
B. intermediate value (i.e. $1mm/yr^{K/4} > b_{cl} > 0.1$ mm/$yr^{K/4}$)

C. high value (i.e. $1mm/yr^{K/4} < b_{cl} < 4mm/yr^{K/4}$)

Note that in the scenario C, the process is unlikely zero-mean stationary. Also, it is mentioned when K is equal to 1 (flicker noise) or 1.5 (power-law noise) in the simulations of the coloured noise.

Figure 1a, 1b and 1c display the results when averaging over 50 time series. The variations are in steps of $[0, 0.3, 0.5, 0.7, 0.8, 1]$ year (see X-axis *[Years]*). We show both the variations of the stochastic and functional models. The Y-axis displays the variations of the models in terms of percentage as discussed in the previous section.

Figure 1. Percentage of variations of the estimated parameters included in the stochastic and functional models when varying the length of the time series. The letters (A), (B) and (C) refer to the various scenarios with different coloured noise amplitudes.



The first result which is common to all three figures, is that the variations in terms of variance of the functional model increases faster than for the results associated with the stochastic model. Previous studies have shown that there is a part of the noise amplitude absorbed in the functional model (Williams , 2003; Montillet et al. , 2015). In our scenario, the estimation of the linear trend may fit partially into the power-law noise, hence reducing the variations of the stochastic noise model. This effect can be amplified with higher spectral indexes (Montillet and Bos , 2019). Now, Figure 1 shows that over 1 year the variations of the stochastic and functional models are less than 4% (on average) for small amplitude coloured noise, whereas when increasing the coloured noise amplitude the variations increase quickly (e.g., more than 20% for the large coloured noise amplitude corresponding to the scenario (C)) . We assume that part of the large variations of the coloured noise is wrongly absorbed in the estimation of the functional model.

Now, Table 2 shows the standard deviation of the difference ($Mean\ Square\ Error$) between the ARMA /FARIMA model and the residuals (i.e. $\mathbf{res}_i$ in Eq. (7)). We do not display any mean, because the fit of the models are done on the zero-mean residuals. Note that the value is averaged over the 50 simulations, together with the variations of the length of the time series described above. The table also displays the averaged correlation between the distribution of the residuals and the Normal or Lévy α-stable distribution. In agreement with the theory, we can see that the ARMA model fits well residuals with small amplitude coloured noise (b_{cl}), whereas with the increase of b_{cl} the FARIMA model fits better than the ARMA

Table 2. Statistics on the error when fitting the ARMA and FARIMA model to the residual time series following the three scenarios.

Error (mm)	K	case A $b_{cl} < 0.1$ mm/$yr^{K/4}$	case B $1mm/yr^{K/4} > b_{cl} > 0.1$ mm/$yr^{K/4}$	case C $1mm/yr^{K/4} < b_{cl} < 3mm/yr^{K/4}$
ARMA	1.0	1.44 ± 0.01	1.74 ± 0.01	1.89 ± 0.04
	1.5	1.46 ± 0.01	1.76 ± 0.04	1.95 ± 0.05
FARIMA	1.0	1.91 ± 0.02	1.85 ± 0.02	1.46 ± 0.02
	1.5	1.89 ± 0.01	1.75 ± 0.03	1.59 ± 0.05

Table 3. Correlation between the distribution of the residuals and the Normal (*Corr. Normal*) distribution or the Lévy α-stable distribution (*Corr. Lévy*) and the Anderson-Darling test (AD) following scenario A, B and C. The results of the AD test is the probability over the 50 trials of accepted null hypothesis. [*Lévy*] or [*Normal*] means the type of distribution uses as the null hypothesis.

Corr. $[0-1]$	K	case A $b_{cl} < 0.1$ mm/$yr^{K/4}$	case B $1mm/yr^{K/4} > b_{cl} > 0.1$ mm/$yr^{K/4}$	case C $1mm/yr^{K/4} < b_{cl} < 3mm/yr^{K/4}$
Corr. Normal	1.0	0.93 ± 0.04	0.92 ± 0.06	0.92 ± 0.04
	1.5	0.92 ± 0.04	0.91 ± 0.04	0.91 ± 0.05
Corr. Lévy	1.0	0.92 ± 0.05	0.94 ± 0.04	0.96 ± 0.03
	1.5	0.93 ± 0.03	0.94 ± 0.03	0.95 ± 0.03
AD test [Normal]	1.0	0.98 ± 0.01	0.96 ± 0.01	0.94 ± 0.03
	1.5	0.97 ± 0.01	0.96 ± 0.02	0.93 ± 0.04
AD test [Lévy]	1.0	0.97 ± 0.02	0.97 ± 0.01	0.97 ± 0.03
	1.5	0.98 ± 0.01	0.97 ± 0.02	0.98 ± 0.02

model. Looking at Table 3 in terms of correlation, the Lévy α-stable distribution fits as good as the Normal distribution as long as the distribution of the residuals is Gaussian without large tails or asymmetry. That is why the Anderson-Darling test accepts the two distributions when the residual time series is Gaussian distributed without tails.In Section 2, we emphasized that the family of Lévy α-stable distributions includes the Normal distribution with specific values for the parameters of the characteristic function (see Eq. (5)). Thus, the results show that for the amplitude of coloured noise corresponding to scenario B (i.e. Intermediate - in Table 2 and 3), the two distributions show similar results. However, the scenario C can potentially generate some aggregation processes in the simulated time series and introducing an asymmetry or large tails in the distribution of the residuals, therefore it emphasizes that the family of Lévy α-stable distributions perform the best in modelling the residuals' distribution. To further support this result, we have added the Anderson-Darling test (Anderson and Darling , 1952) in order to test for the large tails in the distribution of the residuals in Table 3. Note that the results display in Table 2 is the probability of accepting the null hypothesis. Following our previous development, we have used the Normal and the Lévy α-stable distributions as null hypothesis. The results show that this test detects mostly large tails for the scenario C which corresponds to when the family of Lévy α-stable distributions perform better than the Normal distribution.

Finally, those three scenarios support the theory where in the case of small amplitude coloured noise, the stochastic noise properties are dominated by the Gaussian noise, hence defining a third process as a Gaussian Lévy . However, the increase of the coloured noise amplitude shows that it is much more difficult to discriminate between the fractional Lévy and the stable Lévy . The results indicate that the third process can be modelled as a stable Lévy process when mostly the FARIMA fits the residuals due to large amplitude long-memory processes, hence creating a heavy-tail distribution. This result is restrictive for the application to geodetic time series.

3.1.2 Real Time Series

We process the daily position time series of three GNSS stations namely $DRAO$, $ASCO$ and $ALBH$ retrieved from the UNAVCO website (UNAVCO , 2009). The functional model includes the tectonic rate, the first and second harmonic of the seasonal signal, and the occurrence time of the offsets. This occurrence time is obtained from the log file of each station. However, $ALBH$ is known to record slow-slip events from the Cascadia subduction zone (Melbourne et al. , 2005). Thus, we include the offsets provided by the Pacific Northwest Geodetic Array (Miller et al. , 1998). In this scenario we do not know which stochastic model could fit the best the observations. Thus, we use two models: the $PL+WN$ together with the $FN+WN$. Note that in appendices, we display the time series of some of the coordinates, together with the processing and the fitting of the distributions.

Similar to the previous section, Figure 2 displays the percentage of variations of the stochastic and functional models averaged over the East and North coordinates of each station. Note that the average over the three coordinates is displayed in the appendices (see Figure A1). Because the Up coordinate contains much more noise than the East and North coordinates (Williams et al. , 2004; Montillet et al. , 2013), it amplifies the variation of both stochastic and functional models.

Looking at Figure 2, the first result is that for all the stations, there is a strong dependence with the selected noise model. When selecting the power-law noise over the flicker noise model, there is an additional variable to estimate (i.e. the power-law noise exponent, K, in Eq. (4)) within the stochastic noise model. This dependence is already discussed in previous studies (He et al. , 2017, 2019).

The second result is the large variations of the functional model compared with the stochastic model. To recall the simulation results, the functional model partially absorbs the variations of the noise, i.e. the tectonic rate partially fits into the power-law noise. In addition, to some extend at $ASCO$, the sudden increase in the functional model variations at 0.5 year may be explained due to the absorption of some of the noise with the second harmonic of the seasonal signal.

When comparing the variations of the stochastic and functional models with amplitude below 20% for $DRAO$ and $ASCO$, the results agree with the definition of the fractional Lévy process defined in Table 1 as third process modelling the residuals of the East and North components. The variations of the functional model associated with $ALBH$ are much larger than the other two stations, especially for the $PL+WN$ model with variations up to 50%. Those large variations can be explained due to the slow-slip events and the difficulty to model the post-seismic relaxations between two consecutive events He et al. (2019).

Furthermore, Table 4 displays the statistics on the error when fitting the ARMA and FARIMA models to the residuals estimated with the $PL+WN$ stochastic noise model. Figure 3 shows the time series $ASCO$ for the East coordinate using the

Figure 2. Percentage of variations of the estimated parameters included in the stochastic and functional models when varying the length of the daily position GNSS time series corresponding to the stations $DRAO$, $ASCO$ and $ALBH$. The statistics are estimated over the East and North Coordinates.

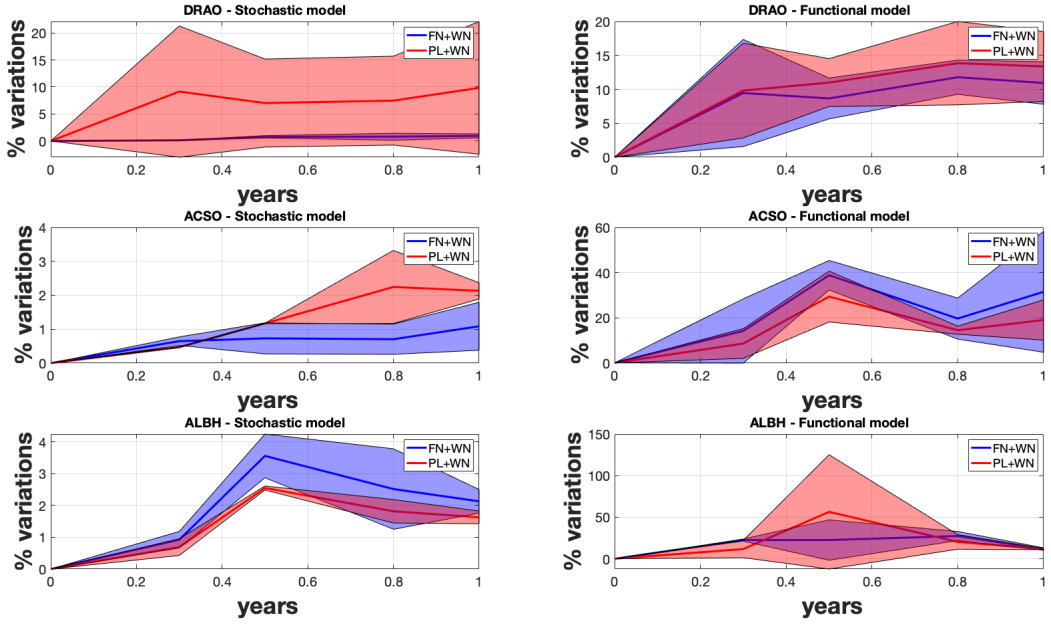


Table 4. Statistics on the Error when fitting the ARMA and FARIMA model to the residual time series for each coordinate of the stations $ALBH$, $DRAO$ and $ASCO$ based on the $PL+WN$ stochastic noise model. Correlation between the distribution of the residuals and the Normal (*Corr. Normal*) and the Lévy α-stable distributions (*Corr. Lévy*). The last column is the Anderson-Darling test. [*Lévy*] or [*Normal*] means the type of distribution uses as the null hypothesis (1 accepted, 0 rejected).

DRAO	*(err. in mm) ARMA*	*(err. in mm) FARIMA*	*Corr. Normal*	*Corr. Lévy*	AD test [*Lévy*]	AD test [*Normal*]
East	1.07 ± 0.01	1.10 ± 0.07	0.94	0.97	1	1
North	1.02 ± 0.02	1.01 ± 0.01	0.96	0.96	1	1
Up	2.32 ± 0.21	2.15 ± 0.30	0.97	0.98	1	1
ASCO						
East	0.77 ± 0.01	0.77 ± 0.06	0.98	0.97	1	1
North	0.84 ± 0.03	0.73 ± 0.03	0.97	0.96	1	1
Up	2.71 ± 0.12	2.34 ± 0.17	0.92	0.96	1	0
ALBH						
East	0.97 ± 0.06	0.87 ± 0.06	0.98	0.98	1	1
North	1.54 ± 0.03	1.06 ± 0.14	0.97	0.98	1	1
Up	4.36 ± 0.17	4.08 ± 0.25	0.92	0.95	1	0

full time series. Note that Table A1 displays in the appendices the results when using the $FN + WN$ stochastic noise model. The FARIMA and ARMA models perform closely for all three stations. The large value for the Up coordinate is due to the amplitude of the noise much larger for this coordinate than for the East and North components. In terms of correlating the distribution of the residuals with the Normal and the Lévy α-stable distribution, the correlation value is relatively the same for all stations which indicates that the distribution of the residuals are Gaussian with the absence of large tails. The Anderson-Darling test also confirms this result when the acceptance of the null hypothesis is the same for the two distributions. Those results further support the selection of the fractional Lévy process as the third stochastic process. However, the study of real time series also underlines the difficulty to characterize statistically this third stochastic process. Note that the Anderson-Darling test shows also that there are some variations for the up coordinate where the Lévy α-stable distribution is only selected. As discussed above, the noise on the up coordinate is much larger than in the other coordinates, therefore it may create small tails.

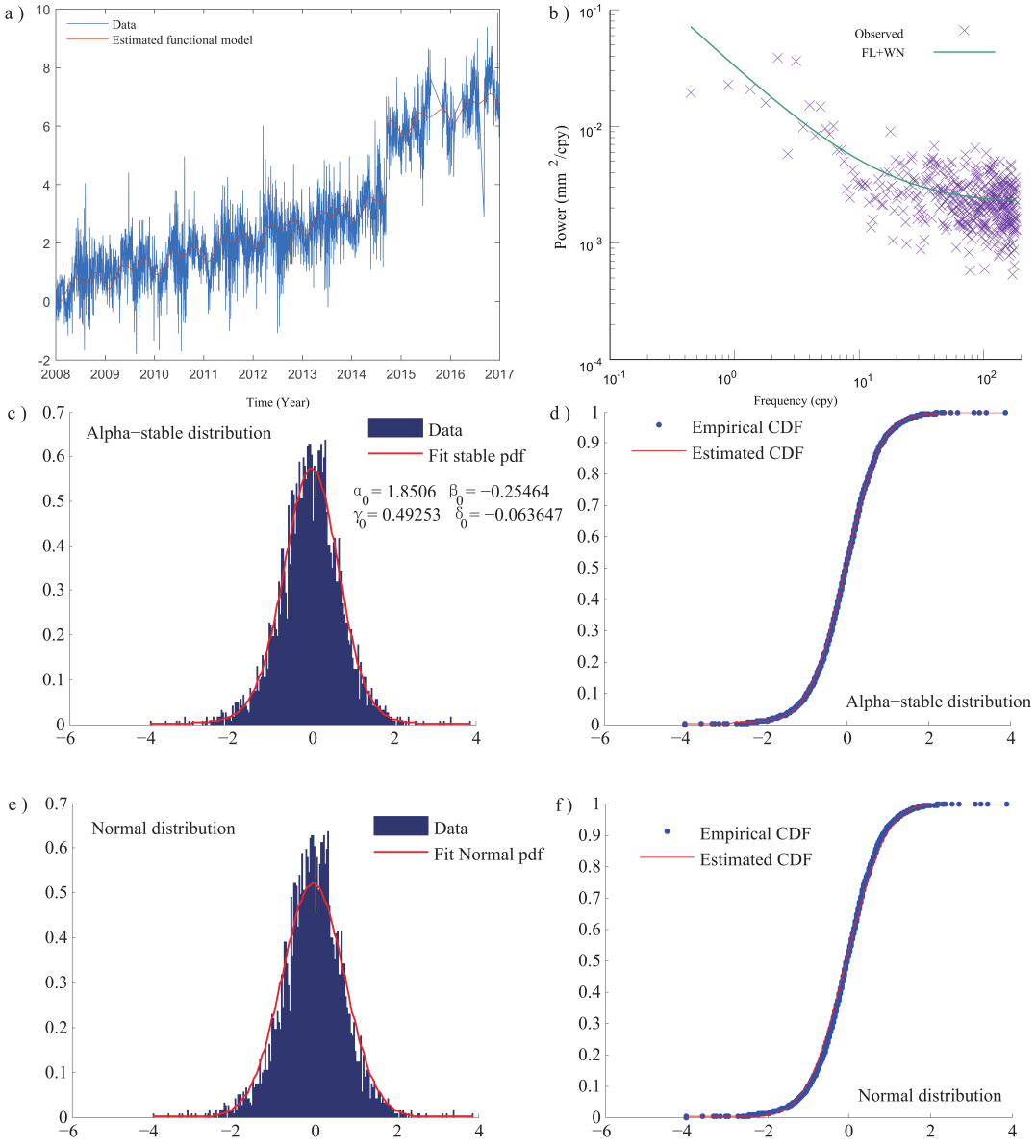
3.2 Discussion on the Limits of Modelling with Lévy Processes

In Montillet and Yu (2015), it was assumed that the infinite variance of the residual time series comes from large tails of the distribution (i.e heavy tails), generated by a large amplitude of coloured noise, outliers and other remaining geophysical signals. The same study implied that the values of the noise variance should be bounded, excluding extreme values. This is an important assumption to decide whether or not (symmetric) Lévy α-stable distributions can be used to model any geodetic time series. This section investigates how the variance due to residual tectonic rate or seasonal signal evolves with the length of the residual time series (i.e. L epochs).

To recall Section 2.1 and the assumption on the noise properties, let us develop the close-form formula of the mean and variance of the residual time series. The residual time series is $\Delta \mathbf{s}_1 = [\Delta s_1(t_1), ..., \Delta s_1(t_L)]$ as defined in Eq. (7). The mean $\langle \Delta s_1(L) \rangle$ and variance $\sigma^2(L)$ are computed over L epochs (i.e. considering the L-th epoch defined as $t_L = Ldt$, with the sampling time dt equal 1 for simplification and without taking into account any data gaps in order to have a continuous time series). Based on Papoulis and Unnikrishna Pillai (2002), one can estimate $\langle \Delta s_1(L) \rangle$ in the cases of a remaining linear trend such as:

$$\begin{aligned} \Delta s_1(t_i) &= a_r t_i + b_r + n(t_i) \\ \langle \Delta s_1(L) \rangle &= \frac{1}{L} \sum_{i=1}^{L} (a_r t_i + b_r + n(t_i)) \\ &= b_r + a_r \frac{(L+1)}{2} + \mu_C \\ &\simeq a_r \frac{L}{2} + \mu_C \end{aligned} \tag{14}$$

Figure 3. GNSS time series for the $ASCO$ station (East coordinate) with the $PL+WN$ model. A/ the time series together with the functional model, B/ the power-spectrum, C/ Residual time series with Lévy α-stable distribution, D/ cumulative density function residual time series and Lévy α-stable distribution (Corr. Lévy $= 0.98$), E/ Residual time series with Normal distribution, F/ cumulative density function residual time series and Normal distribution (corr. Norm. $= 0.97$).



where a_r and b_r are the amplitude and the intersect of the remaining tectonic rate. Note that the subscript r designates *residual* of a geophysical signal in the remaining. $\simeq$ is the approximation for $L \gg 1$. The variance $\sigma^2(L)$ is equal to:

$$\begin{aligned}
\sigma^2(L) &= \frac{1}{L}\sum_{i=1}^{L}(\Delta s_1(t_i) - \langle \Delta s_1(L) \rangle)^2 \\
&= a_r^2 \frac{(L+1)(2L+1)}{6} - a_r^2 \frac{(L+1)^2}{4} + b_r^2 + \frac{2a_r}{L} Cross(a_r, n) + \sigma_n^2(L) - \mu_C(\mu_C + a_r(L+1)) \\
&\simeq \frac{a_r^2 L^2}{12} + \sigma_n^2(L) + b_r^2 - \mu_C a_r L
\end{aligned} \tag{15}$$

Note that $Cross(a_r, n)$ is the cross term between $a_r t_i$ and the noise term $n(t_i)$. Now, if we assume that the remaining seasonal signal $S_r(t)$ is a pseudo periodic function at frequencies similar to the seasonal signal, hence taking the form $S_r(t) = \sum_{j=1}^{N} c_{r,j}\cos(d_j t) + e_{r,j}\sin(d_j t)$. Thus, we can do the same estimation as above in the case of a remaining pseudo periodic component in the residual time series, such as:

$$\begin{aligned}
\Delta s_1(t_i) &= S_r(t_i) + n(t_i) \\
\langle \Delta s_1(L) \rangle &= \frac{1}{L}\sum_{i=1}^{L}(S_r(t_i) + n(t_i)) \\
&\simeq \delta + \mu_C
\end{aligned} \tag{16}$$

where δ is the average of the remaining seasonal signal. It is assumed to be independent of L and bounded such as a periodic function. The variance is equal to:

$$\begin{aligned}
\sigma^2(L) &= \frac{1}{L}\sum_{i=1}^{L}\sum_{j=1}^{N} c_{r,j}^2\cos(d_j t)^2 + e_{r,j}^2\sin(d_j t)^2 + \sigma_n^2(L) \\
&\quad + \frac{2}{L} Cross(S_r, n) - \langle \Delta s_1(L) \rangle^2 \\
&\simeq \sigma_n^2(L) + \sum_{j=1}^{N} c_{r,j}^2 + e_{r,j}^2 - (\delta + \mu_C)^2
\end{aligned} \tag{17}$$

with $Cross(S_r, n)$ is the cross term between $S_r(t)$ and $n(t)$. For all the cross terms, we assume that the deterministic signals and the noise are completely uncorrelated, which is valid with white Gaussian noise (e.g., signal processing - Papoulis and Unnikrishna Pillai (2002)). As previously discussed in Section 2.1, the coloured noise is characterised by long-memory processes, hence producing non-zero covariance with residual signals. Due to the varying amplitude of the coloured noise in geodetic time series with mixed spectra, the uncorrelated assumption is currently debated within the community (Herring et al. , 2016; He et al. , 2017). Therefore, recent works have introduced a random component together with a deterministic signal: nonlinear rate (Wang et al. , 2016; Dmitrieva et al. , 2017), non-deterministic seasonal signal (Davis et al. , 2012; Chen et al. , 2015; Klos et al. , 2018). Thus, strictly speaking, the estimate σ^2 should be seen as an upper bound.

The closed-form solution of the variance $\sigma^2(L)$ shows that the variance is unbounded in the case of a residual linear trend. If this residual trend originates from various sources not well-described in the functional and stochastic model (i.e. undetected jumps, small amplitude random-walk component) of the geodetic time series, the amplitude of this trend should be rather small

($a < 1$ mm/yr) considering the length of GNSS time series available until now ($L < 30$ years). Unless this nonlinear residual trend has a large amplitude, a correction of the functional model must be done a posteriori due to possible anxiety between the models and the observations. The same remarks can be applied to the variance of the remaining seasonal signal where a large amplitude would imply a misfit with the functional model. Thus, we expect rather small amplitude of the coefficients $c_{r,j}$ and $e_{r,j}$ (e.g., $c_{r,j} \sim 0.1$ mm to $e_{r,j} \sim 0.001$ mm). Also, in the appendices, we have developed a similar formula to take into account undetected offsets, where we show that the variance is also bounded. In this case, a large value would mean that one or several large offsets have not been included in the functional model.

4 Conclusions

We have investigated the statistical assumptions behind using the fBm and the family of Lévy α-stable distributions in order to model the stochastic processes within the residual GNSS time series. We model the residual time series as a sum of three stochastic processes. The first two processes are defined from the stochastic model and assumptions on the noise properties of the geodetic time series. The third process is assumed to belong to the Lévy processes. We then distinguish three cases. In the case of a residual time series containing only short-term processes, the process is a Gaussian Lévy process. In the presence of long-term correlations and exhibiting self-similarity property, fractional Lévy processes can be seen as an alternative model of using the fBm. Due to the linear relationship between the Hurst parameter and the fractional parameter of the FARIMA, it is likely that the FARIMA can fit the residual time series under specific conditions (i.e. amplitude of the coloured noise). The third case is the stable Lévy process, with the presence of long-term correlations, high amplitude aggregation processes or random-walk.

In order to check our model, we have simulated mixed spectra time series with various levels of coloured noise. We have then developed a N-step methodology based on varying the length of the time series to study the variations of the stochastic and functional models and also to model the distribution of the residuals. The results emphasize the difficulty to separate the fractional Lévy process and the stable Lévy process mainly due to the absorption of the variations of stochastic processes by the functional model, unless the distribution of the residuals exhibits heavy-tails.

The discussion on the limits of modelling the stochastic properties of the residuals with the stable Lévy process underlines that the infinite variance property can only be satisfied in the case of heavy-tailed distributions, resulting from 1/ the presence of a large amplitude random-walk (e.g., temporal aggregation in financial time series), 2/ an important misfit between the models (i.e. functional and stochastic) and the observations, which means that there is anxiety in the choice of the functional model (e.g., unmodelled large jumps, large outliers). With longer and longer time series, one may be able to statistically characterize more precisely the third stochastic process. Finally, future work should investigate the autoregressive conditional heteroscedasticity (ARCH) model applied to GNSS time series in order to model differently the stochastic properties (e.g. non-stationarity beyond the mean).

Acknowledgements. We would like to thank Dr. Machiel S. Bos from the SEGAL - University of Beira Interior for multiple discussions on the stochastic properties of the GNSS time series. He also develops the Hector software used in this study. Dr. Xiaoxing He would like to acknowledge the Natural Science Foundation of Jiangxi Province supporting his work on GNSS time series (Research on the Noise Model Refinement Algorithm on GNSS Time Series with Non-linear Variation). Finally, we would like to thank the reviewers for the constructive comments helping to improve this manuscript.

Appendix A: Estimation of the Variance in the Presence of Offsets

Here, we model the offsets in the time series as Heaviside step functions according to He et al. (2017). Following Section 3.2, the residual time series in presence of remaining offsets can be written such as:

$$\Delta s_1(t_i) = \sum_{k=1}^{ng} g_k \mathcal{H}(t_i - T_k) + n(t_i) \tag{A.1}$$

Where $\mathcal{H}$ is the Heaviside step function; g_k is the amplitude of the offset; T_k is the time of occurrence of the offset; ng is the number of offsets; n is the noise. One can estimate the average over L epochs:

$$\begin{aligned} \langle \Delta s_1(L) \rangle &= \frac{1}{L} \sum_{i=1}^{L} (\sum_{k=1}^{ng} g_k \mathcal{H}(t_i - T_k)) + \mu_C(t) \\ &= \frac{1}{L} \sum_{k=1}^{ng} g_k \mathcal{H}(t_L - T_k) + \mu_C(t) \end{aligned} \tag{A.2}$$

Note that $\mu_C(t)$ is the mean of the coloured noise, slowly varying in time (see Section 2.1). The variance is equal to:

$$\begin{aligned} \sigma^2(L) &= \frac{1}{L} \sum_{i=1}^{L} (\sum_{k=1}^{ng} g_k \mathcal{H}(t_i - T_k) + n(t_i) - \langle \Delta s_1(L) \rangle)^2 \\ &\simeq \sigma_n^2(L) + \frac{1}{L} (\sum_{k=1}^{ng} g_k \mathcal{H}(t_L - T_k))^2 - (\langle \Delta s_1(L) \rangle)^2 \end{aligned} \tag{A.3}$$

In the presence of small (undetectable) offsets ($g_k < 1$ mm), we can further assume that $\langle \Delta s_1(L) \rangle \sim \mu_C(t)$ and $\sigma^2(L) \sim \sigma_n^2(L) - \mu_C^2(t)$. For multiple large uncorrected offsets (i.e. noticeable above the noise floor), the variance can be large, but the distribution of the residual time series should look like multiple Gaussian distributions overlapping each other corresponding to the segments of the time series defined by those noticeable offsets. This case is not taken into account in our assumptions summarized in Table 1, because it supposes that there is a large anxiety about the chosen functional model (i.e. obviously missing some large noticeable offsets well above the noise floor). Note that for a comprehensive discussion about offset detection, we invite readers to refer to Gazeaux et al. (2013) and He et al. (2017).

Appendix B: Additional Tables and Figures

Table A1. Statistics on the Error when fitting the ARMA and FARIMA model to the residual time series for each coordinate of the stations $ALBH$, $DRAO$ and $ASCO$ based on the $FN + WN$ stochastic noise model. Correlation between the distribution of the residuals and the Normal (*Corr. Normal*) and the Lévy α-stable distributions (*Corr. Lévy*). The last column is the Anderson-Darling test. [*Lévy*] or [*Normal*] means the type of distribution uses as the null hypothesis (1 accepted, 0 rejected).

DRAO	*(err. in mm) ARMA*	*(err. in mm) FARIMA*	*Corr. Normal*	*Corr. Lévy*	AD test [*Lévy*]	AD test [*Normal*]
East	1.07 ± 0.01	1.00 ± 0.02	0.95	0.95	1	1
North	1.02 ± 0.02	1.32 ± 0.07	0.96	0.98	1	1
Up	2.33 ± 0.18	2.20 ± 0.32	0.94	0.96	1	1
ASCO						
East	0.77 ± 0.01	0.75 ± 0.07	0.95	0.96	1	1
North	0.85 ± 0.03	0.74 ± 0.05	0.94	0.96	1	1
Up	2.18 ± 0.14	2.51 ± 0.21	0.93	0.94	1	1
ALBH						
East	0.97 ± 0.04	0.86 ± 0.06	0.95	0.95	1	1
North	1.52 ± 0.08	1.08 ± 0.10	0.96	0.95	1	1
Up	3.83 ± 0.21	3.32 ± 0.15	0.93	0.95	1	0

Figure A1. Percentage of variations of the estimated parameters included in the stochastic and functional models when varying the length of the daily position GNSS time series corresponding to the stations $DRAO$, $ASCO$ and $ALBH$. The statistics are estimated over the East, North and Up Coordinates

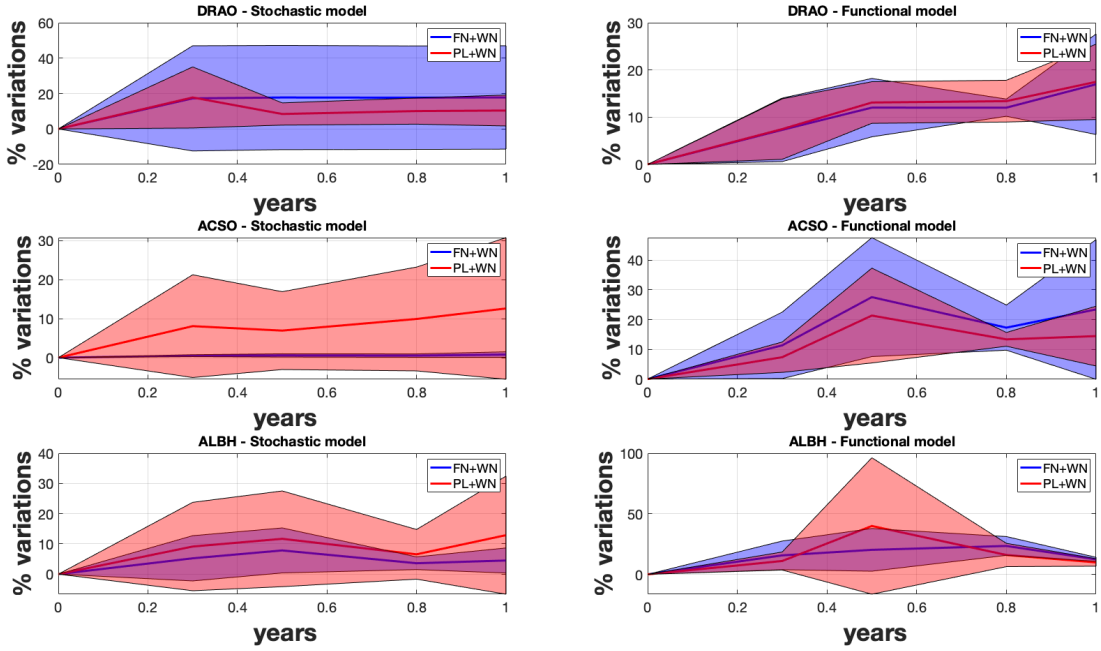


Figure A2. GNSS time series for the $DRAO$ station (North coordinate) with the $FN + WN$ model. A/ the time series together with the functional model, B/ the power-spectrum, C/ Residual time series with Lévy α-stable distribution, D/ cumulative density function residual time series and Lévy α-stable distribution (Corr. Lévy $= 0.98$), E/ Residual time series with Normal distribution, F/ cumulative density function residual time series and Normal distribution (Corr. Norm. $= 0.96$).

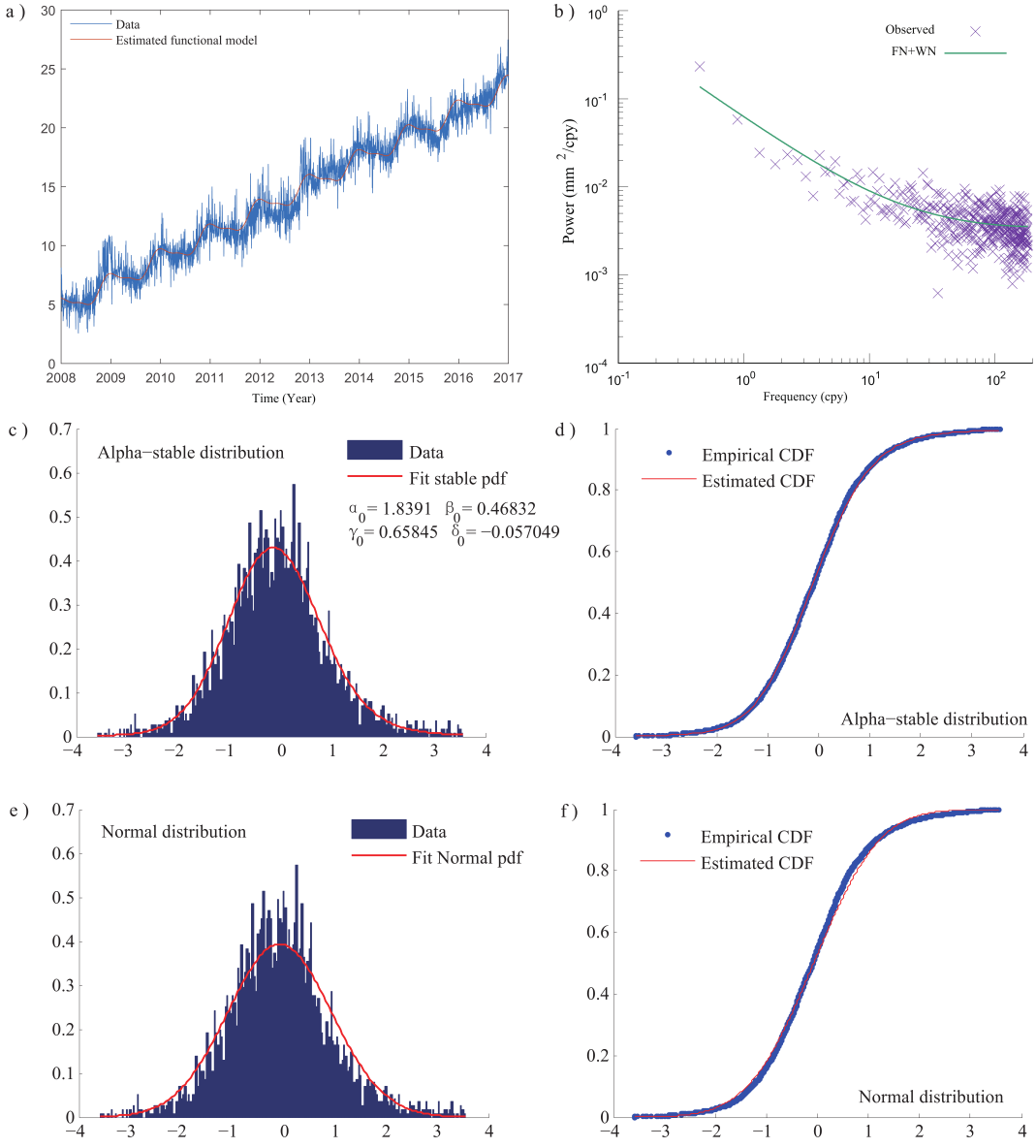


Figure A3. GNSS time series for the $ALBH$ station (East coordinate) with the $PL+WN$ model. A/ the time series together with the functional model, B/ the power-spectrum, C/ Residual time series with Lévy α-stable distribution, D/ cumulative density function residual time series and Lévy α-stable distribution (Corr. Lévy $= 0.98$), E/ Residual time series with Normal distribution, F/ cumulative density function residual time series and Normal distribution (Corr. Norm. $= 0.98$).

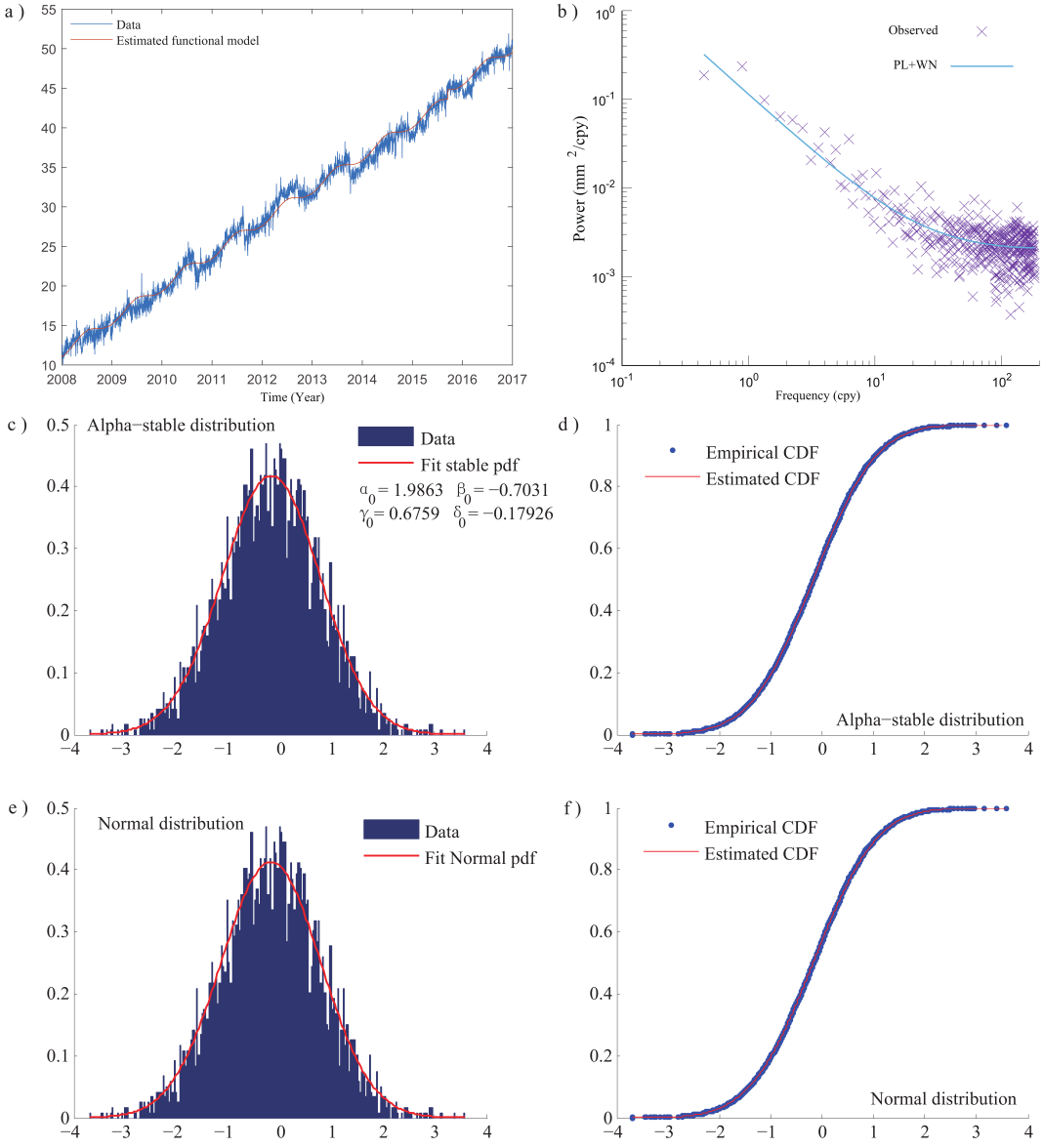
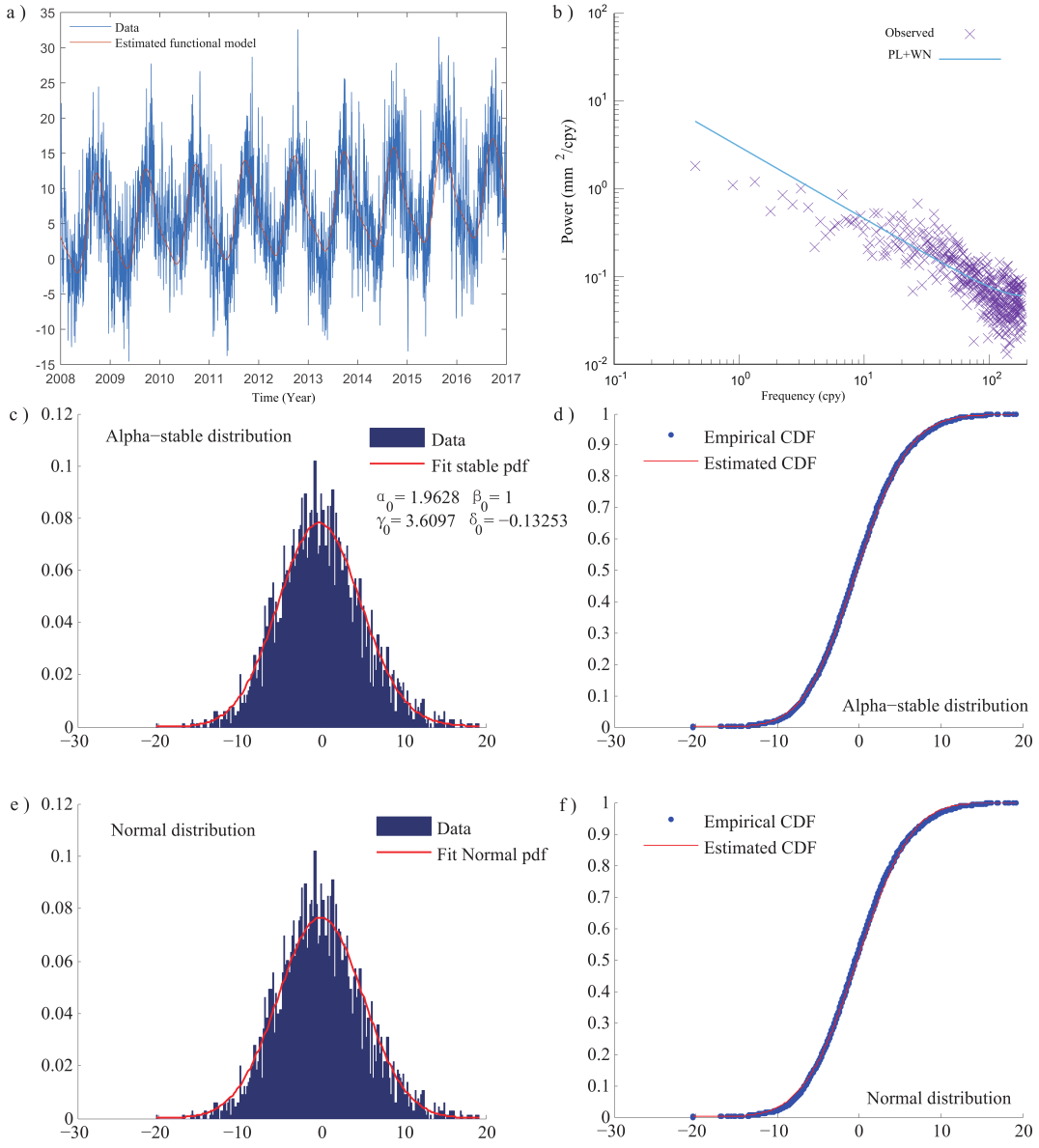


Figure A4. GNSS time series for $DRAO$ (Up coordinate) with the $PL+WN$ model. A/ the time series together with the functional model, B/ the power-spectrum, C/ Residual time series with Lévy α-stable distribution, D/ cumulative density function residual time series and Lévy α-stable distribution(Corr. Lévy $= 0.98$), E/ Residual time series with Normal distribution, F/ cumulative density function residual time series and Normal distribution(Corr. Norm. $= 0.97$).



References

Anderson, T. W., Darling, D. A.: Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. Annals of Mathematical Statistics, 23, p. 193–212, 1952. doi:10.1214/aoms/1177729437

Bevis, M., Brown, A.: Trajectory models and reference frames for crustal motion geodesy. J. of Geod., 88, 283, 2014. doi:10.1007/s00190-013-0685-5.

Bock, Y., Melgar, D.: Physical applications of GPS geodesy: a review. Rep. Prog. Phys., 79 (10), 2016. doi:10.1088/0034-4885/79/10/106801.

Bos, M.S., Fernandes, R.M.S., Williams, S.D.P., Bastos, L.: Fast error analysis of continuous GNSS observations with missing data. J. Geod., 87, 351–360, 2013. doi.org/10.1007/s00190-012-0605-0.

Botai, O.J., Combrinck, L., Sivakumar, V.: Interferences of α-stable distribution of the underlying noise components in geodetic data. South Afr. J. of Geo., 2011. doi:10.2113/gssajg.114.3-4.541.

Chen, Q., Weigelt, M., Sneeuw, N., van Dam, T.: On Time-Variable Seasonal Signals: Comparison of SSA and Kalman Filtering Based Approach, in: Sneeuw N., Novak P., Crespi M., Sanso F. (eds) VIII Hotine-Marussi Symposium on Mathematical Geodesy. International Association of Geodesy Symposia, vol 142. Springer, Cham, 2015.

Davis, J.L., Wernicke, B.P., Tamisiea, M.E.: On seasonal signals in geodetic time series. J. Geophys. Res., 117 (B01403). doi:10.1029/2011JB008690, 2012.

Dmitrieva, K., Segall, P., Bradley, A. M.: Effects of linear trends on estimation of noise in GNSS position time-series, Geophys. J. Int., 208(1), p. 281-288, 2017. doi:10.1093/gji/ggw391.

Eke, A., Herman, P., Kocsis, L., Kozak, L.: Fractal characterization of complexity in temporal physiological signals, Physiol. Meas., 23(1), 2002. doi:10.1088/0967-3334/23/1/201.

Gazeaux, J., Williams, S., King, M., Bos, M., Dach, R., Deo, M., Moore, A. W., Ostini L., Petrie, E., Roggero, M., Teferle, F. N., Olivares, G., Webb, F. H.: Detecting offsets in GPS time series: First results from the detection of offsets in GPS experiment, J. Geophys. Res., 118(5), 2397 - 2407, 2013. doi:doi.org/10.1002/jgrb.50152

Granger, C.W., Joyeux, R. : An introduction to long-memory time series models and fractional differencing, J. Time Ser. Anal., 1, 15-29, 1980. doi:10.1111/j.1467-9892.1980.tb00297.x

Haykin, S.: Adaptive Filter Theory. Fourth edition, Prentice Hall Upper Saddle River, New Jersey, 2002. ISBN-13: 978-0130901262.

He, X., Montillet, J.-P., Fernandes, R.M.S., Bos, M.S., Yu, K., Hua, X., Jiang, W.: Review of current GPS methodologies for producing accurate time series and their error sources. J. Geodyn., 106, 2017. doi:10.1016/j.jog.2017.01.004.

He, X.,Bos, M.S., Montillet, J.-P., Fernandes, R.M.S.: Investigation of information criteria and noise models for GNSS time series. J. Geod., 2018. doi: 10.1007/s00190-019-01244-y.

Herring, T.A., King, R.W., McClusky, S.C.: Introduction to GAMIT/GLOBK, report, MIT, Cambridge, 2010.

Herring, T.A., King, R.W., McClusky, S.C., Floyd, M., Wang, L., Murray, M., Melbourne, T.,Santillan, M., Szeliga, W., Phillips, D., Puskas, C.: Plate Boundary Observatory and Related Networks: GPS Data Analysis Methods and Geodetic Products. Rev. Geophys.,54, 2016. doi:10.1002/2016RG000529.

Klos, A., Bos, M.S., Bogusz, J.: Detecting time-varying seasonal signal in GPS position time series with different noise levels. J. GPS Solut., 22 (21), 2018. doi:10.1007/s10291-017-0686-6.

Koutrouvelis, I.A.: Regression-type estimation of the parameters of stable laws. J. Am. Statist. Assoc., 75, p.918-928, 1980. doi:10.2307/2287182.

Langbein, J.: Noise in GPS displacement measurements from Southern California and Southern Nevada, J. Geophys. Res., 113, B05405, 2008. doi:10.1029/2007JB005247

Langbein, J., Svarc, J. L.: Evaluation of Temporally Correlated Noise in Global Navigation Satellite System Time Series: Geodetic Monument Performance. J. Geophys. Res., 124(1), 925-942, 2019. doi:10.1029/2018JB016783

Li, J., Miyashita, K., Kato, T., Miyazaki, S.: GPS time series modeling by autoregressive moving average method: Application to the crustal deformation in central Japan. Earth Planet Space, 52, 155-162, 2000.

Li, T.-H.: Time Series with Mixed Spectra, CRC Press, 2013. ISBN 9781584881766.

Mandelbrodt, B., Van Ness, J.W.: Fractional Brownian Motions, Fractional Noises and Applications. SIAM Rev., 10(4), p.422-437, 1968.

Melbourne, T. I., Szeliga, W. M., Miller, M., Santillan, V. M.: Extent and duration of the 2003 Cascadia slow earthquake. Geophys. Res. Lett., 32, L04301, 2005. doi:10.1029/2004GL021790.

Miller, M. M., Dragert, H., Endo, E., Freymueller, J. T., Goldfinger, C., Kelsey, H. M., et al.: PANGA: Precise measurements help gauge Pacific Northwest's Earthquake Potential. Eos Transactions, American Geophysical Union, 79(23), 269–275, 1998.

Montillet, J.-P., Tregoning, P., McClusky, S., Yu, K.: Extracting white noise statistics in GPS coordinate time series. IEEE Geosci. Remote Sens. Lett., 10(3):563-567, 2013. doi:10.1109/LGRS.2012.2213576.

Montillet, J.-P., Williams, S.D.P., Koulali, A., McClusky, S.C.: Estimation of offsets in GPS time-series and application to the detection of earthquake deformation in the far-field. Geophys. J. Int., 200(2), 1207-1221, 2015. doi.org/10.1093/gji/ggu473.

Montillet, J.P., Yu, K.: Modeling geodetic processes with levy α-stable distribution and FARIMA. Math. Geosci., 47(6), 2015. doi:10.1007/s11004-014-9574-6.

Montillet, J.-P., and Bos, M.S.: Geodetic Time Series Analysis in Earth Sciences. Springer Geophysics, 2019. doi: 0.1007/978-3-030-21718-1.

Nikias, C.L., Shao, M.: Signal processing with Alpha-Stable Distributions and Applications. New York, Wiley edition, 1995. ISBN-10 : 047110647X.

Nolan, J.P.: Stable Distributions - Models for Heavy Tailed Data. Birkhauser, Boston, 2018. Book online: $http://fs2.american.edu/jpnolan/www/stable/stable.html$.

Panahi, H.: Model Selection Test for the Heavy-Tailed Distributions under Censored Samples with Application in Financial Data. Int. J. of Financial Studiesb(MDPI), 4(4), 2016. doi:10.3390/ijfs4040024.

Panas, E.: Estimating fractal dimension using stable distributions and exploring long memory through ARFIMA models in Athens Stock Exchange, Appl. Fin. Econ., 11(4), p. 395-402, 2001.

Papoulis A., Unnikrishna Pillai S.: Probability, Random Variables and Stochastic Processes. 4th Ed., McGraw-Hill Series in Electrical and Computer Engineering, the McGraw-Hill companies, 2002. ISBN: 0-07-366011-6.

Pipiras, V., Taqqu, M.: Long-Range Dependence and Self-Similarity, (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge: Cambridge University Press, 2017. ISBN:9781139600347.

Samorodnitsky, G., and Taqqu, M. S.: Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance (Chapman and Hall, London), 1994.

Sowell, F.: Modeling long-run behavior with the fractional ARIMA model. J. Monetary Econ., 29, p.277-302, 1991. doi:10.1016/0304-3932(92)90016-U.

UNAVCO: Plate Boundary Observatory: The first five years. Boulder, CO: UNAVCO. Retrieved from $https://www.unavco.org/education/outreach/pamphlets/2009-PBO/PBO-2009-brochure-first-five-years.pdf$, 2009.

Wang, X., Cheng, Y., Wu, S., Zhang, K.: An enhanced singular spectrum analysis method for constructing non-linear model of GPS site movement. J. of Geophys. Res., 121 (3), 2016. doi:10.1002/2015JB012573.

Weron, A., Burnecki, K., Mercik, S., and Weron, K.: Complete description of all self-similar models driven by Levy stable noise. Phys. Rev. E., 2005. doi: 10.1103/PhysRevE.71.016113.

Williams, S.D.P.: The effect of coloured noise on the uncertainties of rates estimated from geodetic time series. J. Geod., 76, p.483-494, 2003. doi:10.1007/s00190-002-0283-4

Williams, S.D.P., Bock, Y., Fang, P., Jamason, P., Nikolaidis, R.M., Prawirodirdjo, L., Miller, M., Johnson, D.J.: Error analysis of continuous GPS position time series. J. Geophys. Res., 109(B03412), 2004. doi:10.1029/2003JB002741.

Wooldridge, J.M.: Econometric Analysis of Cross Section and Panel Data. First edition, MIT Press, 2010. ISBN (13):9780262232586.

Application of Lévy Processes in Modelling (Geodetic) Time Series With Mixed Spectra - Supplementary Material

Jean-Philippe Montillet [1,2], Xiaoxing He [3], Kegen Yu [4], and Changliang Xiong [5]

[1] Physikalisch - Meteorologisches Observatorium Davos / World Radiation Center, Davos, Switzerland
[2]Space and Earth Geodetic Analysis laboratory (SEGAL), University Beira Interior, Covhila, Portugal
[3]School of Civil Engineering and Architecture, East China Jiaotong University, Nan Chang, China
[4]School of Environmental Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, China
[5]Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Science (CAS), Xiaohongshan West Road, Wuhan, China

Correspondence: J.-P. Montillet (jpmontillet@segal.ubi.pt)

A. Geophysical model and ML Estimator

Following Section 2.1 in order to obtain the residual time series $x(t)$ (t the epoch), we subtract the functional model $s_0(t)$ to the GNSS observations $s(t)$ (see Eq. (1)). The functional model of the geophysical signals is based on the polynomial trigonometric method (Li et al. , 2000; Williams , 2003; Tregoning and Watson , 2009).

$$s_0(t) = at + b + \sum_{j=1}^{N} (c_j \cos(d_j t) + e_j \sin(d_j t)) \tag{A.1}$$

with $s_0(t)$ the sum of the tectonic rate (with coefficient a and b) and the seasonal signal (sum of $\cos$ and $\sin$ functions with coefficients c_j and e_j). Note that d_j is equal to $2\pi j/N$, and N can be equal up to 7 (He et al. , 2017). One can also add a Heaviside step function at nominated time t_i in order to estimate the amplitude of an offsets (see appendices). It should be emphasised that the geophysical model is selected based on the surrounding geodynamical activity around the GNSS stations (Montillet and Bos , 2019). Finally, the residual signal is considered to be the remaining geophysical signals (i.e. seasonal component and tectonic rate) not completely estimated due to the mismodelling of the stochastic properties of the time series and other small amplitude (i.e. sub-millimeter) short-time duration transient signals (i.e. local signals, subsidence, ...) (Bos et al. , 2013; Montillet et al. , 2015; Herring et al. , 2016; He et al. , 2017).

Furthermore in this study, the Hector software is used to estimate jointly the functional and stochastic models in order to produce the residual time series as described in Section 2.4. The software is based on a maximum likelihood estimator (MLE). To recall Montillet and Bos (2019) (Chapter 2), for linear models, the log-likelihood for a time series of length L can be rewritten as:

$$\ln(Lo) = -\frac{1}{2}\left[L\ln(2\pi) + \ln(\det(\boldsymbol{C})) + (\boldsymbol{s} - \boldsymbol{A}\boldsymbol{z})^T \boldsymbol{C}^{-1} (\boldsymbol{s} - \boldsymbol{A}\boldsymbol{z})\right] \tag{A.2}$$

This function must be maximised. Assuming that the covariance matrix $\boldsymbol{C}$ is known, then it is a constant and does not influence finding the maximum. Next, the term $(\boldsymbol{s} - \boldsymbol{Az})$ represent the observations minus the fitted model and are called the residual time series $\boldsymbol{x}$. Note that $(\boldsymbol{Az})$ is the matrix notation of $s_0(t)$. The last term can be written as $\boldsymbol{x}^T\boldsymbol{C}^{-1}\boldsymbol{x}$ and it is a quadratic function, weighted by the inverse of matrix $\boldsymbol{C}$.

Now let us compute the derivative of $\ln(Lo)$:

$$\frac{d\ln(Lo)}{d\boldsymbol{z}} = \boldsymbol{A}^T\boldsymbol{C}^{-1}\boldsymbol{s} - \boldsymbol{A}^T\boldsymbol{C}^{-1}\boldsymbol{Az} \tag{A.3}$$

The minimum of $\ln(Lo)$ occurs when this derivative is zero. Thus:

$$\boldsymbol{A}^T\boldsymbol{C}^{-1}\boldsymbol{Az} = \boldsymbol{A}^T\boldsymbol{C}^{-1}\boldsymbol{s} \;\rightarrow\; \boldsymbol{z} = \left(\boldsymbol{A}^T\boldsymbol{C}^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{C}^{-1}\boldsymbol{s} \tag{A.4}$$

This is the weighted least-squares equation to estimate the parameters $\boldsymbol{z}$. Most derivations of this equation focus on the minimisation of the quadratic cost function. However, here we highlight the fact that for observations that contain Gaussian multivariate noise, the weighted least-squares estimator is a maximum likelihood estimator (MLE). Therefore, the Hector software estimates the functional and stochastic parameters via the MLE. Note that in our case $\boldsymbol{C}$ is not a constant, because we assume that the time series contain white and coloured noise. In fact $\boldsymbol{C}$ is equal to the covariance matrix $E\{\mathbf{n}^T\mathbf{n}\}$ in Section 2.1. Thus, the expression of $\boldsymbol{C}$ changes depending on the selection of the stochastic noise model (i.e. Flicker + White noise, Power-law + white noise) discussed in Section 2. Note that further assumptions (i.e. matrix computation) to increase the computational speed can be found in (Bos et al. , 2013) and (Montillet and Bos , 2019). Finally, the Gaussian multivariate noise model (or Gauss-Markov assumption) holds with the additional assumption that the coloured noise is slowly varying which means that 1/ the non-stationarity of the noise around the mean, and 2/ no intermittency in the time series, i.e. no events creating short high bursts or sudden large deviations for the mean.

B. Relationship between the FARIMA, ARMA and fBm

Following Granger and Joyeux (1980), Panas (2001) and Pipiras and Taqqu (2017), a time series (e.g., $\mathbf{x}$) follows an FARIMA (p,d,q) process if it can be defined by

$$\begin{aligned}
\Psi_p(Z)x(t) &= \Theta_k(Z)(1-Z)^{-d}b(t) \\
\Psi_p(Z) &= 1 - a_1 * Z - a_2 * Z^2 - ... - a_p * Z^p \\
\Theta_q(Z) &= 1 + b_1 * Z + b_2 * Z^2 + ... + b_k * Z^q \\
(1-Z)^{-d} &= \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} * Z^j
\end{aligned} \tag{B.5}$$

where $E\{\mathbf{b}\}$ equal zero and $\sigma_b^2 < \infty$. Z is the backshift operator applied to $x\ (t)$ and $b\ (t)$, respectively. The properties of the FARIMA model are presented by Granger and Joyeux (1980): i) if the roots of $\Phi_p(Z)$ and $\Theta_q(Z)$ are outside the unit circle and $d < |0.5|$, then $\mathbf{x}$ is both stationary and invertible; ii) if $0 < d < 0.5$ the FARIMA model is capable of generating stationary

series which are persistent. In this case the process displays long-memory characteristics, with an algebraic autocorrelation decay to zero; iii) if $d \geq 0.5$ the process is non-stationary ; iv) when d equal to 0 it is an ARMA process exhibiting short memory; v) when $-0.5 \leq d < 0$ the FARIMA process is said to exhibit intermediate memory or anti-persistence. This is very similar to the description of the Hurst parameter in the fBm model. Note that one can underline the relationship between d and H such as $H = d + 0.5$, well-known in financial time series analysis in the presence of aggregation processes (Panas , 2001). Note that, the fBm can be defined by its integral representation (see supplementary material - C) or the aggregation of the fractional Gaussian noise (Taqqu et al. , 1995).

C. fBm and fLsm: integral representation and discussion

The fractional Brownian motion (fBm) $\{B_H(t)\}_{t\geq 0}$ has the integral representation:

$$B_H(t) = \int_{-\infty}^{\infty} \left((t-u)_+^{H-\frac{1}{2}} - (-u)_+^{H-\frac{1}{2}}\right) dB(u) \tag{C.6}$$

where $x_+ = max(x, 0)$ and $B(u)$ is a Brownian motion (Bm). It is H-self-similar with stationary increments and it is the only Gaussian process with such properties for $0 < H < 1$ (Samorodnitsky and Taqqu , 1994). It is worth mentioning that a damped version of the fBm exists and known as the Matérn process, defined having a sloped spectrum that matches fBm at high frequencies and taking on a constant value in the vicinity of zero frequency (Lilly et al. , 2017). However, this process is out of the scope of this study.

From Weron et al. (2005), the fractional Lévy stable motion (fLsm) can be defined with the process $\{Z_\alpha^H(t)\}$ (with t in $\mathbb{R}$) by the following integral representation:

$$\begin{aligned} Z_\alpha^H(t) &= \int_{-\infty}^{\infty} \left((t-u)_+^{H-\frac{1}{\alpha}} - (-u)_+^{H-\frac{1}{\alpha}}\right) dZ_\alpha(u) \\ &= \int_{-\infty}^{\infty} f_t(u) dZ_\alpha(u) \end{aligned} \tag{C.7}$$

where $Z_\alpha(u)$ is a symmetric Lévy-stable motion (Lsm). The integral is well defined for $0 < H < 1$ and $0 < \alpha \leq 2$ as a weighted average of the Lévy stable motion $Z_\alpha(u)$ over the infinite past with the weight given by the above integral kernel denoted by $f_t(u)$. The process $Z_\alpha^H(t)$ is H-self-similar and has stationary increments. Note that H-self-similarity follows from the above integral representation and the fact that the kernel $f_t(u)$ is r-self-similar with $r = H - 1/\alpha$, when the integrator $Z_\alpha(u)$ is $1/\alpha$-self-similar. This implies the following important relation:

$$H = r + 1/\alpha \tag{C.8}$$

The representation Eq. (C.7) of fLsm is similar to the representation (C.6) of the fractional Brownian motion. Therefore fLsm reduces to the fractional Brownian motion if one sets $\alpha = 2$. When we put $H = 1/\alpha$ we obtain the Lévy α-stable motion

which is an extension of the Brownian motion to the α-stable case. At the contrary to the Gaussian case ($\alpha = 2$) the Lévy α-stable motion ($0 < \alpha < 2$) is not the only $1/\alpha$-self-similar Lévy α-stable process with stationary increments. This is true for $0 < \alpha < 1$ only (Weron et al. , 2005). Note that this definition of the Fractional Lévy process is different from Benassi et al. (2004), which is not a self-similar process. Finally, another type of processes worth mentioning is the Cauchy-class of processes, which consist of the stationary Gaussian random processes defined by a correlation function which depends on the Hurst parameter which can be seen as the generalization of some stochastic models (Gneiting and Schlather , 2004).

D. Some Considerations on the Simulations of the Mixed Noise in the GNSS time Series

This section recalls one way to simulate the coloured noise in the GNSS time series following Montillet and Bos (2019). Granger and Joyeux (1980) and Hosking (1981) demonstrated that power-law noise can be achieved using fractional differencing of Gaussian noise:

$$(1-B)^{-K/2}\boldsymbol{v} = \boldsymbol{w} \tag{D.9}$$

where B is the backward-shift operator ($Bv_i = v_{i-1}$) and $\boldsymbol{v}$ a vector with independent and identically distributed (IID) Gaussian noise. Hosking and Granger used the parameter d for the fraction $-K/2$ which is more concise when one focuses on the fractional differencing aspect. However, in Geodesy the spectral index K is used in the equations. Hosking's definition of the fractional differencing is:

$$\begin{aligned}(1-B)^{-K/2} &= \sum_{i=0}^{\infty}\binom{-K/2}{i}(-B)^i \\ &= 1 - \frac{K}{2}B - \frac{1}{2}\frac{K}{2}(1-\frac{K}{2})B^2 + \dots \\ &= \sum_{i=0}^{\infty} h_i \end{aligned} \tag{D.10}$$

The coefficients h_i can be viewed as a filter that is applied to the independent white noise. These coefficients can be conveniently computed using the following recurrence relation (Kasdin , 1995):

$$\begin{aligned} &h_0 = 1 \\ &h_i = (i - \frac{K}{2} - 1)\frac{h_{i-1}}{i} \quad \text{for } i > 0 \end{aligned} \tag{D.11}$$

One can see that for increasing i, the fraction $(i - K/2 - 1)/i$ is slightly less than 1. Thus, the coefficients h_i only decrease very slowly to zero. This implies that the current noise value w_i depends on many previous values of $\boldsymbol{v}$. In other words, the noise has a long-memory. Eq. (D.11) also shows that when the spectral index $K = 0$, then all coefficients h_i are zero except for h_0. This implies that there is no temporal correlation between the noise values. One normally assumes that $v_i = 0$ for $i < 0$.

With this assumption, the unit covariance between w_k and w_l with $l > k$ is:

$$C(w_k, w_l) = \sum_{i=0}^{k} h_i h_{i+(l-k)} \tag{D.12}$$

Since $K = 0$ produces an identity matrix, the associated white noise covariance matrix is represented by the unit matrix $\boldsymbol{I}$. The general power-law covariance matrix is represented by matrix $\boldsymbol{J}$. The sum of white and power-law noise can be written as (Williams , 2003).It is a widely used combination of noise models to describe the noise in GNSS time series (Williams et al. , 2004). Besides the parameters of the linear model (i.e. the functional model), maximum likelihood estimation can be used to also estimate the parameters K, σ_{pl} and σ_{wn}. This approach has been implemented various software packages including Hector (Bos et al. , 2013).

We assumed that $v_i = 0$ for $i < 0$ which corresponds to no noise before the first observation. This is an important assumption that has been introduced for a practical reason. For a spectral index K smaller than -1, the noise becomes non-stationary. Most GNSS time series contain flicker noise which is just non-stationary. Using the assumption of zero noise before the first observation, the covariance matrix slowly grows over time but always remains finite.

References

Benassi, A., Cohen, S., and Istas, J.: On roughness indices for fractional fields. Bernoulli, 10(2), 357-373, 2004.

Bos, M.S., Fernandes, R.M.S., Williams, S.D.P., Bastos, L.: Fast error analysis of continuous GNSS observations with missing data. J. Geod., 87, p. 351–360, 2013. doi.org/10.1007/s00190-012-0605-0.

Granger, C.W., Joyeux, R. : An introduction to long-memory time series models and fractional differencing. J. Time Ser. Anal., 1, 15-29, 1980. doi:10.1111/j.1467-9892.1980.tb00297.x

Gneiting, T., Schlather, M.: Stochastic Models That Separate Fractal Dimension and the Hurst Effect. SIAM Review, 46(2), p. 269-282, 2004. doi:10.1137/S0036144501394387.

He, X., Montillet, J.-P., Fernandes, R.M.S., Bos, M.S., Yu, K., Hua, X., Jiang, W.: Review of current GPS methodologies for producing accurate time series and their error sources. J. Geodyn., 106, 2017. doi:10.1016/j.jog.2017.01.004.

Herring, T.A., King, R.W., McClusky, S.C., Floyd, M., Wang, L., Murray, M., Melbourne, T.,Santillan, M., Szeliga, W., Phillips, D., Puskas, C.: Plate Boundary Observatory and Related Networks: GPS Data Analysis Methods and Geodetic Products. Rev. Geophys.,54, 2016. doi:10.1002/2016RG000529.

Hosking, J. R. M.: Fractional differencing. Biometrika, 68, 165–176, 1981. doi:10.1093/biomet/68.1.165.

Kasdin, N.: Discrete simulation of colored noise and stochastic processes and $1/f^{\alpha}$ power-law noise generation. Proc. IEEE, vol. 83, 1995. doi:10.1109/5.381848.

Li, J., Miyashita, K., Kato, T., Miyazaki, S.: GPS time series modeling by autoregressive moving average method: Application to the crustal deformation in central Japan. Earth Planet Space, 52, 155-162, 2000.

Lilly, J.M., Sykulski, A.M., Early, J. J., and Olhede, S. C.: Fractional Brownian motion, the Matérn process, and stochastic modeling of turbulent dispersion. Nonlin. Processes Geophys. 24, 481–514, 2017. doi:10.5194/npg-24-481-2017.

Montillet, J.-P., Williams, S.D.P., Koulali, A., McClusky, S.C.: Estimation of offsets in GPS time-series and application to the detection of earthquake deformation in the far-field. Geophys. J. Int., 200(2), 1207-1221, 2015. doi.org/10.1093/gji/ggu473.

Montillet, J.-P., and Bos, M.S.: Geodetic Time Series Analysis in Earth Sciences. Springer Geophysics, 2019. doi: 0.1007/978-3-030-21718-1.

Panas, E.: Estimating fractal dimension using stable distributions and exploring long memory through ARFIMA models in Athens Stock Exchange. Appl. Fin. Econ., 11(4), p.395-402, 2001.

Pipiras, V., Taqqu, M.: Long-Range Dependence and Self-Similarity, (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge: Cambridge University Press, 2017. ISBN:9781139600347.

Samorodnitsky, G., and Taqqu, M. S.: Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance (Chapman and Hall, London), 1994.

Taqqu, M.S., Teverovsky, V., Willinger, W.: Estimators for long-range dependence: an empirical study. Fractals, 3(4), p. 785-798, 1995. doi: 10.1142/S0218348X95000692.

Tregoning, P., Watson, C.: Atmospheric effects and spurious signals in GPS analyses. J. Geophys. Res., 114 (B09403), 2009. doi:10.1029/2009JB006344.

Weron, A., Burnecki, K., Mercik, S., and Weron, K.: Complete description of all self-similar models driven by Levy stable noise. Phys. Rev. E., 2005. doi: 10.1103/PhysRevE.71.016113.

Williams, S.D.P.: The effect of coloured noise on the uncertainties of rates estimated from geodetic time series. J. Geod., 76, p.483-494, 2003. doi:10.1007/s00190-002-0283-4

Williams, S.D.P., Bock, Y., Fang, P., Jamason, P., Nikolaidis, R.M., Prawirodirdjo, L., Miller, M., Johnson, D.J.: Error analysis of continuous GPS position time series. J. Geophys. Res., 109(B03412), 2004. doi:10.1029/2003JB002741.

Rebuttal letter for the manuscript: Application of Lévy Processes in Modelling (Geodetic) Time Series with Mixed Spectra

Dear Editor, reviewers,

Thank you for your comments. We have modified the manuscript based on the comments of the reviewers. During the discussion phase, we did reply to the major comments and we have included some suggestions that we have followed for this revised manuscript.

The major modifications are:

- Section 2.1 has been rewritten based on R1 comments
- The N-steps algorithm is described with more details.
- The simulations of the GNSS time series have been revised to include the Anderson-Darling test as suggested by R2
- We have also improved the Figures via simulations and also reduced the steps in the variation of the length of the time series (which also improved the Figures with real time series).
- We have also applied the Anderson-Darling test to the real time series
- As R2 pointed out, we have put some of the appendices in the "supplementary material". Thus, Appendix A, B, C and E are treated as supplementary material (as A, B, C, D). Only D and F should remain as appendices (as Appendix A, Appendix B). The supplementary material is dissociated from the manuscript to comply with the NPG guidelines.

Finally, R2 has pointed out an interesting research direction with the ARCH/GARCH model which will be the topic of our next work.

Thank you for your comments.

[RC1/SC1]

Major comments

- *« The mathematical model is really messily written, and it is hard to understand it. Take all the definitions from the appendix, and put them in the text, and remove all thespeculative/descriptive material. »*

  (R). As underlined in the discussion (SC1), there are a few points which must be emphasized when working with geodetic data. Thus, we have rewritten the whole Section 2.1 to include the comments from R1. Also, we reduced the descriptive material in this section.
  Some more details are now in the Supplementary material A.

- *"After you have done the basic model, you can construct an a posteriori distribution of the unknown parameters, and then the target is to sample these parameters withMCMC methods or obtain optimisation-based MAP/ML estimators. Define explicitly your estimators with respect to posterior distribution. Of course you can use some other constructions as well, but define your estimators explicitly. The N-step method is totally heuristic and should not be included in the manuscript. Please come up with some mathematics/stats-based parameter estimation algorithm."*

(R) We have discussed our vision in the SC1 comprehensively. However, the idea of the reviewer is similar to what R2 pointed out for future work, on investigating the use of the ARCH model. We believe that is valuable and interesting approach for a future research. We have taken this in the revised manuscript (see the conclusions).

Note that the N-step method is further explained. It emphasises that we use the ML estimator (as highlighted in section 2.1) to estimate jointly the stochastic and functional model.

Minor Comments

l.27 long-memory processes > deleted – see sentence

l. 43 the sentence now reads: “The Levy stable distributions can model the heavy tail characteristics of some data sets with generally infinite variance.”

p.3 Eq. 3
Section 2.1 including Eq. (1) is rewritten completely. See modifications.
Also, we consider the noise “n” as a continuous process, with mean equal to the coloured noise (as the white noise is zero mean by definition).

p.3
The covariance of the coloured noise “J” is named following the literature (see e.g., Williams et al. 2004 [eq. 4], Bos et al. Chapter 2 [eq. 27] in Montillet and Bos 2019).

p.3 “"Note that the length of the geodetic time series (L) considered in this study is at least 9 years (3285 observations)." Can you plot this data”

Several examples of time series are given with the real time series in Section 3.1.2 and Appendix B (additional figures)

p.3 " …which states that the noise is Gaussian distributed, therefore n follows ..” > corrected – now it reads " …which states that the noise is Gaussian distributed.

p.3 L82 “p. 3: Line 82 – You have not defined β, and for all β <0, the matrix J(β)does not exist. Be more precise.”

This section has been revised, where β is now K in [0, 2]. Note that the definition of K with the different values is given after Eq. 4.

Thank you for your nice comments.

[RC2/AC1]

Major comments

1/ *“Can you explain a bit more systematically how (and why) different types of non-stationarity are associated to the different components of your stochastic model?”*

(R): According to the discussion (RC2), we have added the sentence (l.73-74): “The coloured noise results from various parameters during the processing of the GNSS observations such as the mismodelling of GNSS satellites orbits, … ”

2/ *About the length of the time series*

(R) we modified the text according to the RC2 (l 93-94): “Note that the value of L is here at least 9 years (3285 observations) in order to model correctly the long-range dependencies associated with the coloured noise and to detect slow transient signals …”

3/ *“I am wondering about non-stationarity beyond just the mean, i.e., possible time-dependence inthe variance (or even higher-order distributional characteristics), which I believe can-not be captured by the present model setup but would require either some multiplicative component (e.g., functional model times noise of a certain type)or a stochastic model component beyond ARMA/FARIMA that inherets the property of conditional heteroscedasticity ”*

(R) We discuss it in the RC2. We added some more information in the sentences (Supplementary material A – last sentence) “Finally, the Gaussian multivariate noise model (or Gauss-Markov assumption - see Section 2.1) holds with the additional assumption that the coloured noise is slowly varying which means that 1/ the non-stationarity of the noise around the mean, and 2/ no intermittency in the time series, i.e. no events creating short high bursts or sudden large deviations for the mean.”

- *About the use of Levy Flights*

As discussed in the RC2, it is difficult to include this model within GNSS time series, due to the assumption that GNSS time series have a finite variance and the assumptions that the small amplitude transient signals should not generate jumps larger than the noise floor. Therefore, we have left this model out for now.

- *Intermittency*

Following the discussion in RC2, it is mentioned in the revised manuscript in the assumptions related to the Gauss-Markov model (see added sentence above and in Section 2.1)

- *Hector Software:*

It is only mentioned once in the main text. The use of the software to decrease the log-likelihood function (via MLE) is described in the supplementary material (A). We also added an acknowledgment.

- *Homoscedasticity vs Heteroscedasticity and the ARCH/GRACH model*

(R) We have discussed in the RC2 that our model is based on homoscedasticity with the sum of two (three) stochastic processes. However, we would like to circumvent a discussion on this property. We believe that it will be more appropriate on a future work about the GARCH model. We have added at the end of the conclusions: “Finally, future work should investigate the autoregressive conditional heteroscedasticity (ARCH) model applied to GNSS time series in order to model differently the stochastic properties (e.g. non-stationarity beyond the mean).”

- *The Anderson-Darling Test:*

(R) About the results in Table 3 and 4 (and A.1) using the correlation coefficient. As mentioned in the header of this letter, we have implemented the Anderson-Darling test. Therefore, the results section is revised accordingly. It supports the previous results where the Levy alpha stable distribution mostly detect high tails for the last scenario (c).

Minor comments:

- We change “soil motion” by “ground motion”

- The justification of the length of the time series has been added as discussed above.

- *“ Can you explain the relationship between slowly varying mean of colored noise andthe Gauss-Markov assumption a bit more explicitly? ”*

  (R) According to the discussion, we revised the paragraph L95-100.

- *About the Cauchy-class of processes and H*

  In addition of the discussion in RC2, we have added in the supplementary material (C) : “Finally, another type of processes worth mentioning is the Cauchy-class of processes, which consist of the stationary Gaussian random processes defined by a correlation function which depends on the Hurst parameter which can be seen as the generalization of some stochastic models … ”

- *In the context of this brief discussion of FARIMA models …*

  (R) The relationship between FARIMA and H is discussed in the supplementary material (B)

- *The definition includes a variable k that I don’t find appearing anywhere before.*

  (R) In the revised manuscript, “k” is now “beta”, which is the symmetry parameter in eq. (5)

- *About L154-155:*
  As discussed in the RC2, we have modified the sentence such as:
  “Therefore, the residual time series withholds some remaining unmodelled geophysical signals or unfiltered large outliers which can potentially undermine the Gauss-Markov assumption (e.g., presence of heavy tails in the distribution of the residual time series) ”

- *L.171: In which sense do you consider colored noise to be non-stationary*

  (R) as discussed in RC2, the coloured noise is here non-stationary around the mean. This sentence is revised.

- *L.243: The list of values given in the text is inconsistent with that shown in the figure*.

  (R) there is a confusion with the figures in the previous version. Now “K” is the power-law exponent and it is the same in the figures and the text.

- *L.246: I would not speak of “earlier” here due to the low number of data points in the figure, but rather refer to the overall values of the variance.*

  (R) The sentence is revised accordingly.

- *L. 262 “the driving parameters “*

  (R) replaced with “the parameters of the characteristic function”

- *L.271-274: I understand this as that heavy tails in the series can either be attributed to the residuals or to the third component. Can such an attribution be actually unique?*

(R) As discussed in RC2, it is not clear what the reviewer means. It seems the question is clearly given. Why not understand?

- *L.325 and several times later: There are quite a few cases of equations spanning …*

  (R) We have corrected this issue, eliminating the right hand-side.

- *L.348: "colored noise can generate long-memory processes" seems a bit odd to me*

  (R) we revised the sentence: "coloured noise is characterised by long-memory processes"

- *L.349: Doesn't the mentioned varying amplitude of the colored noise rather call formultiplicative/heteroscedastic models?*

  (R) We value the idea of testing heteroscedastic models such as ARCH and GARCH for the future work. As previously said, this study is mainly based on a sum of various noise components. We have added a sentence in the conclusions.

- L.373: *"long-term correlations"* > corrected

- L.434: instead of the term "hyperbolic", "algebraic" seems more commonly used " > corrected

- *Ll.417-430: Please emphasize that the expressions in Z denote composite operators in terms of the backshift operator Z applied to x(t) and b(t), respectively* > corrected

- *L.436-439: The link to fBm is a bit unclear here. Since you consider stationary pro-cesses (FARIMA class), it would be more reasonable in my opinion to link this to fractional Gaussian noise (fGn) the aggregation of which than provides sample paths for fBm.*

  (R) We added "Note that, the fBm can be defined by its integral representation (see supplementary material - C) or the aggregation of the fractional Gaussian noise … "

- *L.451: extend the equation by an expression including f_t(u) as the latter is used in the text below*

  (R) We added the kernel f_t(u) in Eq. (C.7) in the supplementary material.

  Finally, we have followed your editorial recommendations by checking the languages and moving some of the appendices into the supplementary material.