Stochastic synthesis approximating any process dependence and distribution

Panayiotis Dimitriadis* and Demetris Koutsoyiannis

Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Heroon Polytechneiou 5, 15880 Zographou, Greece

*corresponding author, email: pandim@itia.ntua.gr

Abstract

An extension of the symmetric-moving-average (SMA) scheme is presented for stochastic synthesis of a stationary process by approximating any dependence structure and marginal distribution. The extended SMA model can exactly preserve an arbitrary second-order structure as well as the high order moments of a process, thus enabling a better approximation of any type of dependence (through the second-order statistics) and marginal distribution function (through statistical moments), respectively. Interestingly, by explicitly preserving the coefficient of kurtosis, it can also simulate certain aspects of intermittency, often characterizing the geophysical processes. Several applications with alternative hypothetical marginal distributions, as well as with real world processes, such as precipitation, wind speed and grid-turbulence, highlight the scheme's wide range of applicability in stochastic generation and Monte-Carlo analysis. Particular emphasis is given on turbulence, in an attempt to simulate in a simple way several of its characteristics regarded as puzzles.

Keywords: stochastic modelling; grid-turbulence; hourly surface wind speed; daily precipitation; intermittency; Kumaraswamy distribution; normal-inverse-Gaussian distribution

1. Introduction

The scientific interest on stochastic dynamics has increased over the last decades as an alternative approach of deterministic or chaotic models to simulate the so-called random (i.e. unexplained or unpredictable) fluctuations recorded in non-linear physical processes. Koutsoyiannis (2010) argues that distinguishing deterministic from random is a false dichotomy. Randomness can emerge even in a fully deterministic system with non-linear dynamics. The dice throw, regarded as the emblem of randomness, is such an example (Dimitriadis et al., 2016b). The line distinguishing whether determinism (i.e. predictability) or randomness (i.e. unpredictability) dominates is related to the scale (or length) $l(\varepsilon)$ of the time-window within which the future state deviates from a deterministic prediction by an error threshold ε. For errors smaller than ε, we assume that the system is predictable within a time-window $l(\varepsilon)$ and for larger errors unpredictable. Therefore, by applying stochastic analysis we identify the observed unpredictable fluctuations of the system under investigation with the variability of a devised stochastic process. This stochastic process enables generation of an ensemble of realizations, while observation of the given natural system can only produce a single observed time series (or multiple ones in repeatable experiments). The simplest and yet powerful technique to reveal and analyze entirely the system's variability, is the Monte-Carlo approach. However, this technique requires a generation algorithm capable of modelling any selected marginal probability distribution and second-order dependence structure of the stochastic processes, appropriate for the investigated natural system.

Although there are several methods for simulating arbitrary stochastic process, they all have limitations and advantages (Lavergnat, 2016 and references therein). For example, the non-linear method for the preservation of the distribution function (i.e., a Gaussian distributed process with the desired dependence structure is produced and then transformed to the desired distribution through a non-linear transformation or the inverse distribution function) is often applied for synthesis of long-term (cyclostationary) processes (e.g., Koutsoyiannis et al., 2008), but it has a disadvantage of e.g., distorting the dependence structure (because of the transformation), while, in addition, the transformation cannot be invariant with respect to the time scale (Lombardo et al. 2012). A generalization of this implicit scheme is the so-called copula in which the joint distribution function is approximated, albeit through numerical methods. The latter implicit approaches, whose detailed overview is beyond the scope of this paper (the interested reader is referred to Hoeffding, 1940; Frechet, 1951; Sklar, 1959; as well as to more recent publications, e.g. by Nelsen, 2006 and references therein). Typically they apply non-linear transformations to the processes, often based on the autocovariance function (see Appendix D). However, it is widely known that despite the high flexibility of copulas, the fractal and Hurst-Kolmogorov behaviour (i.e. strong correlation structures at zero scale and long-term persistence at infinite scale, respectively) cannot be easily handled (e.g. Lavergnat, 2016; Ibragimov and Lentzas, 2017). The reason for this, is that the process structure is invariant at these scales (as they tend to zero and infinity) and thus, the structure of the originally generated process is preserved rather than the back-transformed one (see also Appendix D for an example).

Another category of existing methods concern the preservation of the dependence structure through autoregressive models such as the SAR model (Sum of many AR(1) or ARMA(1,1) models; Koutsoyiannis, 2002; Dimitriadis and Koutsoyiannis, 2015a, supplementary material sect. 3). Although these models can simulate a variety of dependence structures, they have a disadvantage if preservation of high order moments is of interest (see also sect. 3.3.1 in Dimitriadis, 2017). A rigorous and general method is the symmetric-moving-average (SMA) scheme introduced by Koutsoyiannis (2000), further advanced by Koutsoyiannis (2016) and implemented within the Castalia computer package (Efstratiadis et al., 2014). This method can fully preserve in an exact way any second-order structure of a process and, simultaneously, the complete multivariate distribution function if it is Gaussian (because of the preservation of the Gaussian attribute within linear transformations). Koutsoyiannis (2000) also studied the application of the same scheme to non-Gaussian processes by preserving the skewness of the marginal distribution. Here, we extend the SMA scheme so that it preserves exactly four (or, if necessary, more) central moments of the distribution, while simultaneously preserving in an exact way any type of second-order dependence structure, such as short-range (e.g., Markov) or long-range (e.g., Hurst-Kolmogorov, abbreviated as HK). Note that the term 'HK behaviour' corresponds to the behaviour of process at large scales while the process itself could not be necessarily an HK process or follow a Gaussian distribution. For example, both the fractional Gaussian noise (fGn) and the generalized-HK (GHK; see below) process are processes exhibiting an HK behaviour, but while the former's autocorrelation function is a power-law type at the whole range of scales, the latter's autocorrelation function is a power-law type only at large scales (at small scales behaves like a Markov process) and its distribution function is not necessarily Gaussian.

To our knowledge there is no other method that can preserve explicitly (i.e. fully analytical calculations) four (or more) marginal moments of a process for any type of dependence structure. In most problems preservation of four moments suffices for a very good approximation of the distribution function. In particular, the fourth moment has been regarded of great importance in some problems, e.g., in the characterization of intermittency in turbulence (Batchelor and Townsend, 1949). Therefore, turbulence is an ideal field for application of the proposed framework. This study has given particular emphasis on turbulence, which is included as the final and most detailed of the case studies presented, in an attempt to show that several aspects of turbulence that are regarded as puzzles can be easily reproduced by a simple model without a major effort. While all applications presented below handle moments up to fourth, the methodology proposed can handle explicitly even higher moments (see Appendix A) and thus even approximate the joint (univariate, multivariate) structures that extend beyond the second-order statistics. Specifically, higher order moments for different time scales can also be adequately approximated, since this scheme can explicitly preserve the high-order moments of the lowest scale, as well as those of the largest scales (which by virtue of the Central Limit Theorem are expected to be close to the Gaussian distribution and thus, can be precisely represented by the second-order dependence structure; for an example see section 4.4). The only limitation of this methodology is that the marginal distribution is approximated to a desired degree, rather than precisely preserved (particularly in non-divisible distributions). This limitation may create difficulties in variables with upper or lower bounds, since these can be only treated in an ad-hoc manner (see Section 4.2). However, this limitation rarely concern practical applications to geophysical processes.

In section 2, we present in detail the computational scheme for preserving in an exact way the dependence structure. We also explain why the preservation of solely the second-order joint statistics can be adequate for capturing the most important attributes of a physical process and suggest that often it is impractical to estimate high-order statistics from observations. In section 3, we present the generation algorithm for simultaneously preserving an approximation of the marginal probability function (through cautiously selected distributions as described in Appendices A and B), a task that we deal with when the actual process distribution departs from normality, along with the dependence structure of the process. Finally, in section 4 we apply the generation scheme to various examples in order to highlight not only its robustness but also its use as a statistical tool to investigate the stochastic nature of complex geophysical processes such as surface wind speed, daily precipitation and turbulent phenomena dominated by intermittent behaviour.

2. Generation scheme for preserving the dependence structure of a process

In this section, we discuss the practical limitations of using popular multi-parameter stochastic models for geophysical processes. In contrast, the SMA generation scheme (Koutsoyiannis, 2000; 2016) is very convenient in preserving the stochastic structure of a process based only on a parsimonious representation of the second-order statistics. Important advantages of this scheme are the parsimony of parameters and the fact that it can deal with non-Gaussian distribution, as will be detailed in the next session.

2.1. Background and notation on stochastic processes

As introduced by Kolmogorov (1931, 1933), a stochastic process $\underline{x}(t)$ is a family of infinitely many random variables, commonly (but not exclusively) indexed by time t, here assumed to be continuous. A random variable $\underline{x}$ is an abstract mathematical object that can take on all of its possible values according to a (marginal) distribution function $F(x)$. In addition to the marginal distribution, a stochastic process is also characterized by its dependence structure. Note that in the above notation we are using the Dutch convention, where an underlined symbol denotes a random variable. A random variable $\underline{x}$ should not be confused by its realizations (e.g., observations) x and a stochastic process $\underline{x}(t)$ should not be confused with its realizations (sample functions or time series) $x(t)$. A realization is usually known (e.g. by observation) only in discrete time, at time intervals of length D, either by taking observations (sampling, with approximately zero response time) at equidistant times iD (i = 1, 2, …) or by averaging the process at each interval D (see more details in Koutsoyiannis, 2016). In the former case, the sampling operation at equidistant times evokes to introduce the discrete-time stochastic process $\underline{x}_i := \underline{x}(iD)$. In the latter case, the observations are not of the process $\underline{x}(t)$ per se, but of the time-averaged process:

$$\underline{x}_i^{(D)} = \frac{1}{D}\int_{(i-1)D}^{iD} \underline{x}(\xi)\mathrm{d}\xi \qquad (1)$$

where i denotes discrete time. While in the definition of the discrete-time process i takes on infinite values ($i \in \mathbb{N}$), in an observed time series it obviously has an upper limit, the total number of observations n, determined from the observation period $T \geq D$, i.e. $n = \lfloor T/D \rfloor$. In the analysis below, to avoid confusion, we will omit the superscript (D) in the notation of the process and we will use $\underline{x}_i$ regardless of the manner the discrete-time process is constructed from the continuous-time one. We also assume that $\underline{x}(t)$ is a stationary and ergodic process (and hence $\underline{x}_i$ too). Note that the proposed SMA model can be used for simulation of both stationary and non-stationary processes (by extracting the deterministic model and transform the original non-stationary process to stationary) following the methodology described in Appendix C. The marginal characteristics of the process are estimated through the classical central moments and the dependence structure of the process is estimated through $\gamma(k)$, i.e. the variance of the averaged process $\underline{x}_i^{(k)}$ vs. scale k, here called climacogram (also, hastily, referred to as aggregate variance), where $k = \kappa D$ is the continuous-time scale in time units and κ the dimensionless discrete one, assuming that D is a time unit that is used for discretization (see Eqn. 1). The climacogram is directly linked to the autocovariance $c(h)$ by $c(h) = 1/2\mathrm{d}^2\big(h^2\gamma(h)\big)/\mathrm{d}h^2$ (e.g. Koutsoyiannis, 2016), where h is the continuous-time lag (in time units). Also, the autocovariance is linked to the power spectrum by $s(w) = 4\int_0^\infty c(h)\cos(2\pi wh)\,\mathrm{d}h$, where w is the frequency (in inverse time units). Thus, each of these three stochastic tools contains exactly the same information. However, it has been shown that the climacogram provides better estimates than the other two (Dimitriadis and Koutsoyiannis, 2015a) and therefore, all applications here are based on that.

2.2. The impracticality of stochastic modelling with many parameters

Several families of autoregressive models are typically used in stochastic modelling with the most popular in literature being model families so-named AR, ARMA, ARIMA, FARIMA. These models are easy to handle and fast in stochastic generation once their parameters are known and not too many. However, whenever the process exhibits long-range dependence these models require a large number of parameters (except only in the FARIMA(0,*d*,0) special case) to preserve in an exact way up to a large number of autocorrelation estimates. Conversely, the typically available observation records can support the estimation of only a few parameters (Lombardo et al. 2014; Koutsoyiannis, 2016). Interestingly, most geophysical processes exhibit such long-term behaviours, as expected considering the maximization of entropy (Koutsoyiannis, 2011) and as verified in several geophysical processes (O'Connell et al., 2016) and specifically in key hydrometeorological processes such as: solar radiation and wind speed (Tsekouras and Koutsoyiannis, 2014; Koutsoyiannis, et al., 2018); precipitation (Iliopoulou et al., 2016); paleoclimatic temperature reconstructions (Markonis and Koutsoyiannis, 2013); and temperature, dew point, wind, precipitation and pressure processes classified by the Köppen-Geiger scheme (Dimitriadis et al., 2017). Additionally, the more complicated dependence structures and similarities identified e.g. among the microscale of turbulent, wind and precipitation processes (Dimitriadis et al., 2016a and references therein) increase the need for parsimonious stochastic approaches.

An additional difficulty may arise when estimating the prediction intervals of a long-range process (Papoulis, 1990, pp. 240-242; Tyralis et al., 2013). Even if the model parameters are calculated with adequate accuracy, this does not guarantee an adequate approximation of the prediction intervals (e.g., see sect. 3.3.1 in Dimitriadis, 2017). Finally, the above model families may confront difficulties even in case of short-range processes but with a non-Gaussian marginal distribution. Only in case of the AR(1) model can non-Gaussian distributions and/or seasonality be simulated simultaneously (e.g., the PGAR model of Fernandez and Salas, 1986) while in higher order autoregressive models this is not possible. For example, consider the simplest model of the ARMA family, which is the ARMA(1,1) model:

$$\underline{x}_i = a\underline{x}_{i-1} + \underline{v}_i + b\underline{v}_{i-1} \tag{2}$$

where $\underline{v}_i$ is a white noise process and a, b are parameters. This model can simulate well short-range dependence, e.g., a Markov process in continuous time discretized by averaging in time steps $D > 0$ (Dimitriadis and Koutsoyiannis, 2015a), but it cannot handle explicitly marginal moments beyond the second moment (Koutsoyiannis, 2016). If simulation of moments higher than the second order is needed, then in the parameter estimation equations mixed joint moments of the form $\mathrm{E}\left[\underline{x}_{i-1}^{p}\underline{v}_{i-1}^{q}\right]$ will emerge (with p, $q > 0$, $p + q > 2$), which are not possible to directly estimate from observed data (since $\underline{v}$ is artificial white noise) or handle (because $\underline{x}_{i-1}$ is not independent from $\underline{v}_{i-1}$). A good alternative could be not to use such models but rather a sum of multiple AR(1) models (SAR), to approximate the correlation structure up to the desired lag by the sum of many independent AR(1) models, with their coefficients theoretically derived rather than arbitrarily calculated. This was introduced and applied for 3 AR(1) models by Koutsoyiannis (2002) and further developed for an arbitrary large number of components by

Dimitriadis and Koutsoyiannis (2015a). However, all of these models (AR, ARMA, SAR etc.) still have several limitations, as for example, it is impossible for them to preserve some important stochastic properties as scale tends to zero in continuous time (see section 4). Interestingly, the SMA model overcomes all the aforementioned limitations as also shown in the next sections and in the applications of this paper.

2.3. Derivation of the SMA coefficients

In the SMA scheme, the simulated process is expressed through the sum of products of coefficients (not parameters) a_j and white noise terms $\underline{v}_i$, i.e. (Koutsoyiannis, 2000):

$$\underline{x}_i = \sum_{j=-l}^{l} a_{|j|} \underline{v}_{i+j} \tag{3}$$

where l theoretically equals infinity but a finite number can be used for preserving the dependence structure up to lag l (Koutsoyiannis, 2016). Also, for simplicity and without loss of generality we assume that $\mathrm{E}[\underline{x}] = \mathrm{E}[\underline{v}] = 0$ and $\mathrm{E}[\underline{v}^2] = \mathrm{Var}[\underline{v}] = 1$. This scheme can be used for stochastic generation of any type of second-order process structure represented by functions such as the climacogram, the autocovariance function, the power spectrum or the variogram. It exhibits several advantages over widely used backward moving average (BMA) schemes, the most important being that it allows closed expressions for the coefficients a_j, based on any of the above functions (Koutsoyiannis, 2000).

As an example, let us consider the HK process, whose climacogram in continuous time is:

$$\gamma(\kappa) = \frac{\gamma(D)}{\kappa^{2-2H}} \tag{4}$$

where $\kappa = k/D$ denotes discrete time scale, $\gamma(D)$ is the variance at the unit time scale D, and H is the Hurst parameter ($0 < H < 1$).

For an HK process with $H > 0.5$ the SMA coefficients can be estimated analytically (Koutsoyiannis, 2016):

$$a_j = C\left(\frac{|j+1|^{H+\frac{1}{2}} + |j-1|^{H+\frac{1}{2}}}{2} - |j|^{H+\frac{1}{2}}\right) \tag{5}$$

where

$$C = \sqrt{\frac{2\Gamma(2H+1)\sin(\pi H)\gamma(D)}{\Gamma^2(H+1/2)\left(1+\sin(\pi H)\right)}} \tag{6}$$

whereas $\Gamma(x)$ is the gamma function.

Another example that will be used below is the so-called Hybrid Hurst-Kolmogorov (HHK) process (Koutsoyiannis 2016), whose climacogram is:

$$\gamma(k) = \frac{\lambda}{(1 + (k/q)^{2M})^{\frac{1-H}{M}}} \tag{7}$$

where λ is the variance of the continuous-time process $\underline{x}(t)$, M is a fractal parameter (see also Gneiting 2000; Gneiting and Schlather 2004; Gneiting et al. 2012; Dimitriadis 2017 and references therein), H is the Hurst parameter and q is a characteristic time parameter. A particular case of the HHK, which will be later used and referred to as GHK, is when M = ½, i.e.:

$$\gamma(k) = \frac{\lambda}{(1 + k/q)^{2-2H}} \tag{8}$$

An explicit expression for the coefficients a_j for the HHK or the GHK may not be easy to derive. However, they can be numerically calculated through the Fourier transform of the discrete power spectrum of the coefficients which is directly linked to the discrete power spectrum of the process (Koutsoyiannis, 2000):

$$s_{\mathrm{d}}^{a}(\omega) = \sqrt{2 s_{\mathrm{d}}(\omega)} \tag{9}$$

where $s_{\mathrm{d}}^{a}(\omega)$ and $s_{\mathrm{d}}(\omega)$ are the power spectra of the SMA coefficients and of the discrete time process, respectively, and ω=D/k is the dimensionless frequency.

3. Generation scheme for approximate preservation of the marginal distribution of a process

In this section, we first discuss the natural and statistical intrinsic limitations of fitting multi-parameter probability functions to geophysical processes, while we emphasize the need to fit and reproduce non-Gaussian distributions, which very often appear in geophysical variables. A non-Gaussian distribution can still be parsimonious in terms of parameters (e.g., two-parameter marginal distributions are often used). While fitting low-parametric marginal distributions to a certain variable is easy, stochastic generation schemes can hardly deal with non-Gaussian distributions and can hardly handle moments higher than second-order. Here, we introduce an extension of the SMA generation scheme for approximating the marginal probability function of a process by exactly preserving its first four central moments which is found to be adequate for various distributions commonly applied in geophysical processes, while in Appendix A we extend the method for even higher moments. We emphasize that these moments that are to be preserved are not necessarily estimated from data. On the contrary, in typical sizes their estimation from data is strongly discouraged as explained below. Rather, the values of these moments should be obtained theoretically, once a specific distribution function is fitted.

3.1. The impracticality of estimating high-order moments in geophysical processes

Non-Gaussian distributions are very common in Nature. It has been shown (Lombardo et al., 2014) that the estimation of high order raw moments is highly uncertain and, thus, it is inefficient to use the schemes described in the previous section to preserve high moments for natural processes with limited sample sizes, as is the case for typical geophysical records. In the case of a continuous HK process the variance of the mean estimator is $\gamma(D)/n^{2-2H}$ (e.g. Koutsoyiannis, 2003), where n is the sample size. Consequently, for estimating the population mean μ of a process with a standard error $\pm\varepsilon$, we would require a time series of length at least $(\sigma/\varepsilon)^{1/(1-H)}$, where $\sigma = \sqrt{\gamma(D)}$ is the standard deviation at scale D (Fig. 1). For example, for an HK process with $H = 0.8$, to estimate the mean of the process with an error $\varepsilon \approx \pm 10\%\sigma$, we need a time series of length $n = 10^5$. Such lengths are hardly available in observations of geophysical processes, which are not only characterized by HK behaviour (a conclusion based on the maximization of entropy as derived in Koutsoyiannis, 2011) but also include sub-daily and seasonal periodicities (e.g., Dimitriadis and Koutsoyiannis, 2015b) that complicate the estimation further.

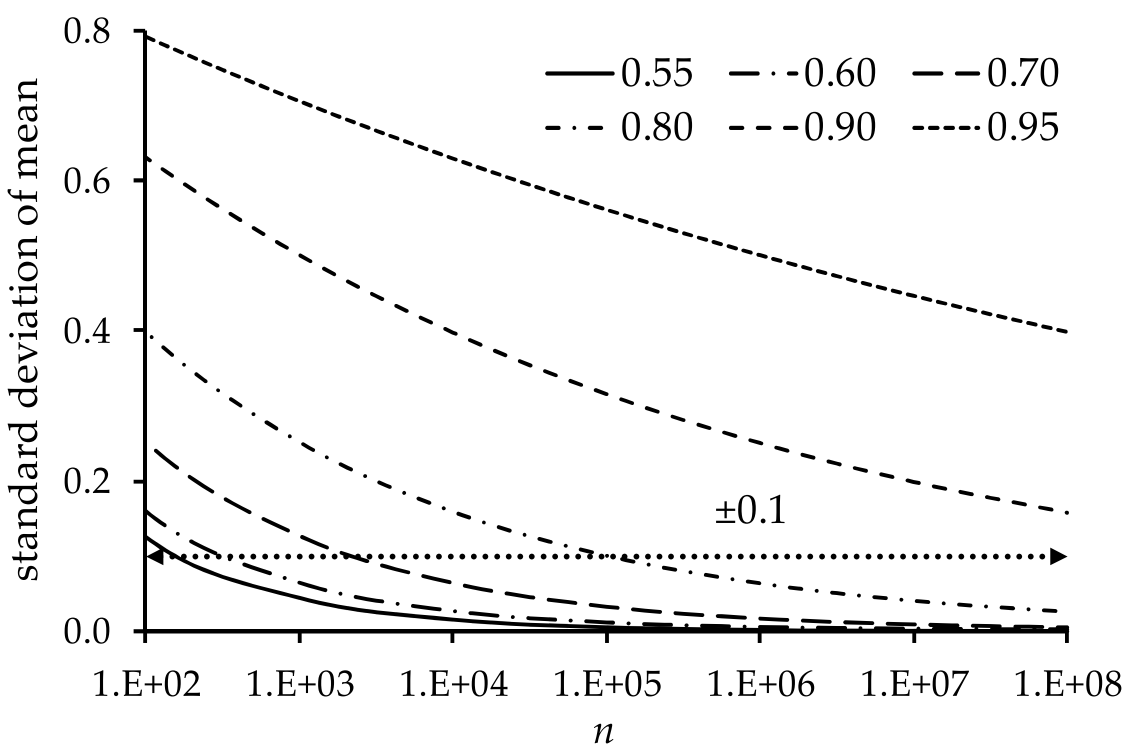

Figure 1: Standard deviation of the mean estimator of an HK process standardized by σ vs. the sample size (n) for various Hurst parameters.

To give another example, we perform a Monte Carlo experiment for an HK process with $H = 0.8$ that follows a standard Gaussian distribution (i.e. $\mu = 0$ and $\sigma = 1$) and the results are shown in Fig. 2. For each synthetic time series we estimate the mean, standard deviation as well as skewness and kurtosis coefficients for six different lengths, i.e., $n = 10$, 10^2,..., 10^6. This experiment shows that for $n = 10^6$ the uncertainty (measured in terms of the standard deviation of sample estimates of each property) is below 10% for all these moments. Therefore, to adequately estimate these moments from data we would need time series with at least similar lengths.

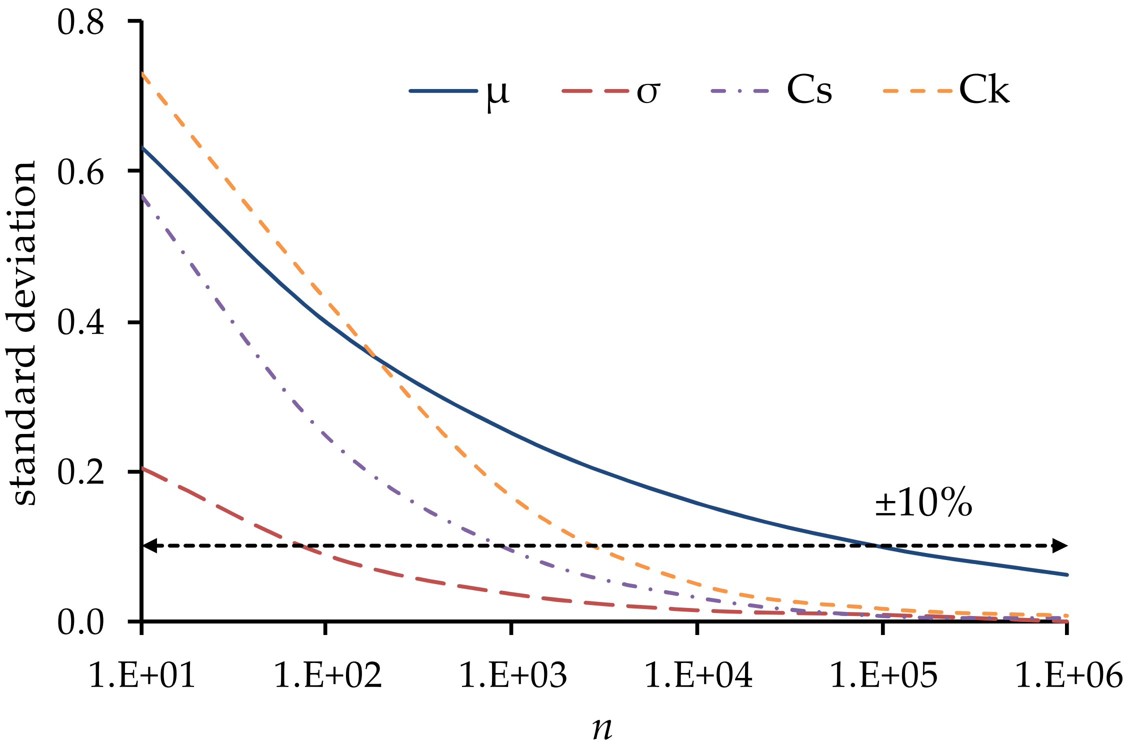

Figure 2: Standard deviation of the sample estimates of the mean (μ), standard deviation (σ), skewness coefficient (C_s) and kurtosis coefficient (C_k) of an HK process with H = 0.8 and N(0,1) distribution vs. the simulation length.

3.2. Derivation of the SMA distribution parameters

The SMA method can explicitly preserve high order marginal moments. However, as already mentioned above, high-order moments cannot be estimated reliably from data while non-Gaussianity can be easily verified empirically but also derived by theoretical reasoning (Koutsoyiannis 2005; 2014). A simple way to simulate non-Gaussian distributions is to calculate theoretically (not from the data but rather from the distribution model) their moments and then explicitly preserve these moments in simulation. Koutsoyiannis (2000) estimated the first three moments of the marginal distribution of the white noise process $\underline{v}_i$, required to reproduce those of the actual process $\underline{x}_i$, using the SMA scheme. With the conventions used in this paper (see above), the mean and variance of $\underline{v}_i$ are 0 and 1, respectively, while the third moment, which is equal to the coefficient of skewness, is:

$$C_{s,v} = \frac{\left(\sum_{j=-l}^{l} a_{|j|}^2\right)^{3/2}}{\sum_{j=-l}^{l} a_{|j|}^3} C_{s,x} \tag{10}$$

where $C_{s,x}$ is the coefficient of skewness of $\underline{x}_i$.

Although preservation of three central moments usually provides acceptable approximations to the theoretical distributions, a non-Gaussian distribution cannot be precisely preserved. Here, we expand the calculations to include the coefficient of kurtosis of the white noise (see Appendix A for proof):

$$C_{\mathrm{k},v} = \frac{\left(\sum_{j=-l}^{l} a_{|j|}^{2}\right)^{2}}{\sum_{j=-l}^{l} a_{|j|}^{4}} C_{\mathrm{k},x} - 6\frac{\sum_{j=-l}^{l-1} \sum_{k=j+1}^{l} a_{|j|}^{2}\, a_{|k|}^{2}}{\sum_{j=-l}^{l} a_{|j|}^{4}} \tag{11}$$

where $C_{\mathrm{k},\underline{x}}$ is the coefficient of kurtosis of $\underline{x}_i$. Note that the constant term in the right-hand side depends only on the SMA coefficients and not on the marginal distribution of the process. Also, note that the kurtosis of the white noise is not proportional to the kurtosis of the process, as is in the case of the skewness (Eqn. 10).

For the generation of the auxiliary variable $\underline{v}$ we need distributions that: (a) contain at least four parameters, creating in such way a large variety of combinations between the first four moments; (b) have closed analytical expressions for the first four central moments; and (c) can easily and quickly generate random numbers. Here, we propose the four-parameter Kumaraswamy (1980) distribution, which is mostly appropriate for generating thin-tailed distributions, and the four-parameter normal inverse Gaussian distribution (e.g., Barndorff-Nielsen, 1978) for generating heavy-tailed distributions. The details of the distributions are contained in Appendix B.

4. Applications

In this section, we present several applications of the extended SMA scheme, first to non-Gaussian white noise process with several two-parameter marginal distributions used extensively in geophysics and we show that even complicated distributions can be well approximated by their first four central moments. We then apply the scheme to a 130-year daily precipitation time series, by fitting an HK model along with a three-parameter marginal distribution. Also, we apply the scheme to multiple hourly wind speed time series recorded over a wide area, by fitting a GHK model along with a three-parameter marginal distribution. Finally, we apply the scheme into a massive database of experimental time series of turbulent velocities recorded at high frequency, by fitting an HHK process and by approximating the unknown marginal distribution with the first four empirical moments.

4.1. Application to white noise processes with various two-parameter distributions

In the examples below, we apply the extended generation scheme to Weibull, gamma, Pareto and lognormal distributions and we illustrate that the preservation up to the fourth moment is adequate for capturing the main body of the distribution as well as a part of the tail (see Fig. 3, which also displays the parameter sets of the specified distributions). As all these variables used in our examples are non negative, any generated negative values are set to zero. This does not cause any distortion worth discussing, as the approximation of the probability density function by the four moments is satisfactory (see Fig. 3 and other figures below), and since this density is typically zero in negative values, the number of generated negative values is negligibly small.

We expect that using the approximation based on the first four moments, the distribution which we actually simulate is the maximized entropy (ME) distribution produced by constraining these moments. The ME resulting distribution, i.e. $f(x;\boldsymbol{\lambda}'):=\mathrm{e}^{\lambda_0{}'+\lambda_1{}'x+\lambda_2{}'x^2+\lambda_3{}'x^3+\lambda_4{}'x^4}$ (Jaynes, 1957), can also be written as:

$$f(x;\boldsymbol{\lambda}):=\frac{1}{\lambda_0}\mathrm{e}^{-\left(\frac{x}{\lambda_1}+\mathrm{sign}(\lambda_2)\left(\frac{x}{\lambda_2}\right)^2+\left(\frac{x}{\lambda_3}\right)^3+\left(\frac{x}{\lambda_4}\right)^4\right)} \tag{12}$$

where $\boldsymbol{\lambda}=[\lambda_0,\lambda_1,\lambda_2,\lambda_3,\lambda_4]$, with λ_0, $\lambda_1,\lambda_2,\lambda_3,\lambda_4$ (with $\lambda_4\ \geq 0$) having same units as x and with constraints:

$$\int_{-\infty}^{\infty}x^r f(x;\boldsymbol{\lambda})\mathrm{d}x=\mathrm{E}\left[\underline{x}^r\right],\ \text{for}\ r=0,\ldots,4 \tag{13}$$

The solution to the above system of equations for $\boldsymbol{\lambda}$ can be achieved through optimization or other numerical algorithms (cf. Balestrino et al., 2009). Parameter $\lambda_0>0$ equals $1/f(0;\boldsymbol{\lambda})$ (where the denominator is the value of the ME probability density at $x=0$). For the estimation of the ME distribution parameters we minimize the error ε_f defined in Eqn. (14), which is based on the absolute value of the difference between the main body of the empirical and modelled distribution along with their left and right tails:

$$\varepsilon_f=\sum_i\left|1-\frac{f_\mathrm{m}(x_i)}{f_\mathrm{e}(x_i)}\right|\sum_i|f_\mathrm{e}(x_i)-f_\mathrm{m}(x_i)|\sum_i\left|1-\frac{f_\mathrm{e}(x_i)}{f_\mathrm{m}(x_i)}\right| \tag{14}$$

where f_m and f_e are the model and empirical distribution functions, respectively.

The quantities $1/\lambda_1, 1/\lambda_2, 1/\lambda_3, 1/\lambda_4$ can be also regarded as weighting factors representing the dependence of the distribution on each raw moment. Interestingly, after standardizing these four parameters based on the sum of their absolute values, $1/\lambda_1$ contributes to the Weibull, gamma, lognormal and Pareto distributions of Fig. 3, approximately 65%, 66%, 69% and 93%, respectively. Similarly, the contribution of $1/\lambda_2$ is approximately 20%, 20%, 18% and 4%, the contribution of $1/\lambda_3$, 11%, 10%, 9% and 2% and the contribution of $1/\lambda_4$, 4%, 4%, 4% and 1%, respectively. Therefore, we may use the ME probability density to approximately determine the weight for each statistical moment. As a rough indicator of the goodness of fit, the correlation coefficient between the theoretical values of these four distributions and the simulated ones is estimated as 99.57%, 99.38%, 99.26%, and 99.84%, respectively.

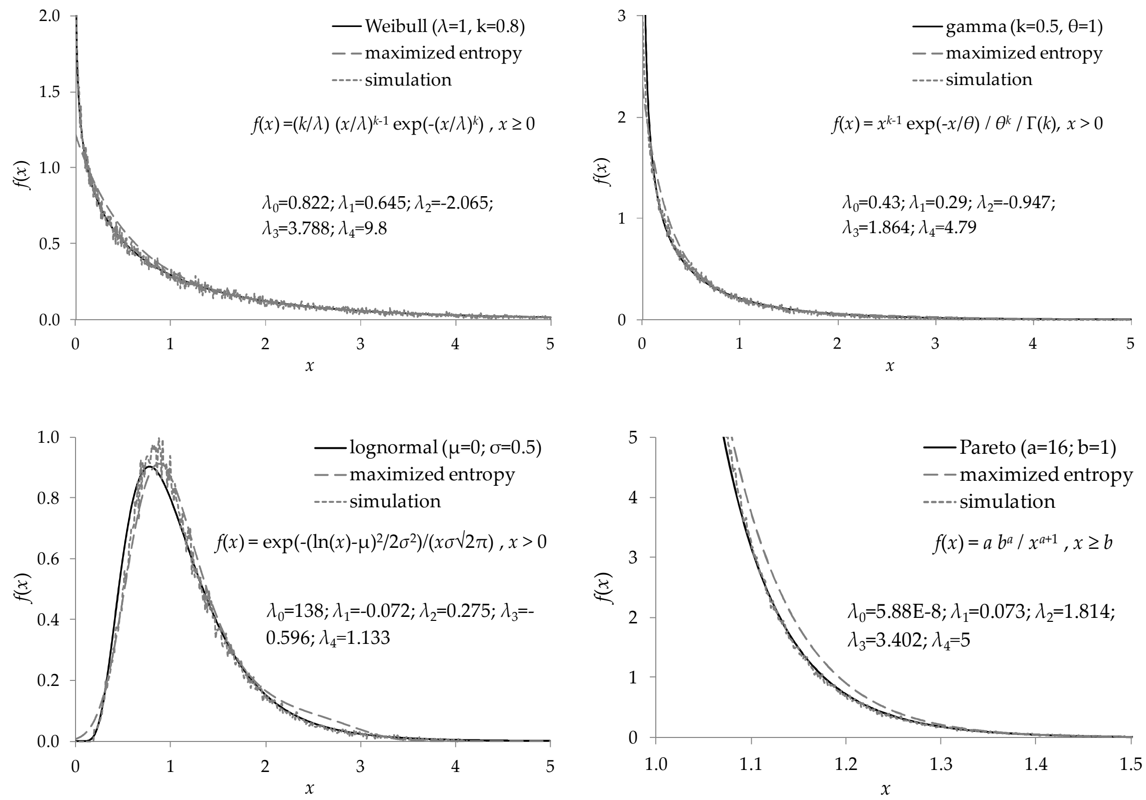

Figure 3: Various two-parameter distributions along with the fitted ME probability density function and the empirical probability density from one single simulation with n = 10^5 using the proposed generation scheme.

4.2. Application to daily precipitation; an HK process with a three-parameter distribution

In this application, we analyse one of the longest daily precipitation time series recorded for over 100 years at the site of Hohenpeißenberg in Germany (latitude 47.801°N, longitude 11.011°E; data from www.gkd.bayern.de/). We apply an HK process (Eqn. 4) with a single continuous-state Pareto II marginal distribution (a special case of the Pareto-Burr-Fuller—PBF distribution, i.e. $F(x) = 1 - (1 + (x/a - h)^b\ )^{-c}$; Koutsoyiannis et al., 2018), introduced for use in precipitation by Koutsoyiannis (2004a) and theoretically justified by Koutsoyiannis (2004b):

$$F(r) = 1 - \left(1 + \left(\frac{r}{a} - h\right)\right)^{-c} \tag{15}$$

where $r > ah$ is precipitation; $a > 0$ is a dimensionless scale parameter; $c > 0$ is a dimensionless parameter characterizing the right tail (extreme events) of the distribution and h is a dimensionless parameter representing a threshold value and characterizing the left tail (dry events) of the distribution.

Theoretically h = 0, but values slightly different from zero not only highly improve fitting (Fig. 4), but also preserve the left tail of the distribution (i.e. probability dry), by simulating the probability dry through $F(0) := \mathrm{P}(\underline{r} \leq 0)$. An h different from 0 is also physically justified since precipitation measurements are usually corrupted with significant uncertainties (Krajewski et al., 1998; Villarini et al., 2008) causing losses mostly due to wind effects (Nespor and Sevruk, 1998) and so, a slightly larger amount of precipitation is expected to be lost before measured. After the generation we can set to zero any negative values of the synthetic time series in order to emulate the observed distribution function. This approach of mixing wet and dry events within a single distribution function is rather simple but can sometimes provide good results (Fig. 4; see also Dimitriadis , 2017, sect. 6, where a more generalized distribution performs an even better fit). For a more accurate approach, in terms of the simulation of the wet/dry probability, one could separate these events and model their joint distribution instead (Lombardo et al., 2017, and references therein).

To account for the seasonal periodicity of precipitation (Langousis and Koutsoyiannis, 2006) we apply a non-linear transformation based on the known marginal distribution of the process which is also preserved. Note that here, the process exhibits weak seasonality that only causes a small increase in the dependence structure, as depicted in the intermediate area of the climacogram in Fig. 4. Therefore, for the simulation of the seasonality we may then use the inverse non-linear transformation (i.e. concept of homogenization, Dimitriadis, 2017, sect. 2.1) or apply a cyclostationary model, where each cycle is treated separately but with the same white noise process, thus, additionally preserving the cross-correlations values (Dimitriadis, 2017, sect. 3.3.3).

To account for estimation bias, since we have a single time series, we apply the innovative method of estimating the parameters of the dependence structure of the process through the mode second-order measure (e.g. climacogram, autocovariance, power spectrum etc.) rather than the expected one (see also Dimitriadis et al., 2016c). For this, we apply a Monte-Carlo analysis by generating one thousand daily time series of 130 years following the fitted marginal distribution and an HK process. From the Monte-Carlo ensemble, we calculate the mode for each scale with an acceptable accuracy and construct the mode climacogram for the specified process. For the estimation of the parameters of the marginal distribution we minimize the same norm as in Eqn. (14) and for the parameters of the dependence structure we use a similar one:

$$\varepsilon_\gamma = \sum_\kappa \left|1 - \frac{\gamma_\mathrm{m}(\kappa)}{\gamma_\mathrm{e}(\kappa)}\right| \sum_i |\gamma_\mathrm{e}(\kappa) - \gamma_\mathrm{m}(\kappa)| \sum_i \left|1 - \frac{\gamma_\mathrm{e}(\kappa)}{\gamma_\mathrm{m}(\kappa)}\right| \tag{16}$$

where γ_m is the model (i.e., the mode for this application) climacogram and γ_e is the empirical climacogram estimated from the classical estimator:

$$\underline{\hat{\gamma}}(\kappa D) = \frac{1}{\lfloor n/\kappa \rfloor - 1} \sum_{i=1}^{\lfloor n/\kappa \rfloor} \left(\underline{\overline{x}_i} - \underline{\overline{x}}\right)^2 \tag{17}$$

where $\lfloor n/\kappa \rfloor$ is the integer part of n/κ, $\bar{\underline{x}}_\kappa = \left(\sum_{l=\kappa(i-1)+1}^{\kappa i} \underline{x}_l\right)/\kappa$ is the sample average of the time-averaged process $\underline{x}_\kappa$ at scale $\kappa = k/D$ as defined in Eqn. (1) and $\bar{\underline{x}} = \sum_{l=1}^{n} \underline{x}_l /n$ is the sample average at scale $\kappa = 1$.

The parameters of the process are estimated as $a = 42.25$ mm, $c = 7.7$, $h = -0.1$, $\gamma(D) = r_{\mathrm{s}}{}^2$, with $r_{\mathrm{s}} = 6.5$ mm the standard deviation of the process, and $H = 0.6$. Through a single synthetic time series of equivalent length and after setting negative values to zero, the modelled marginal characteristics are estimated as: $\mu = 3.3$ (3.1) mm, $\sigma = 6.5$ (6.5) mm, $C_{\mathrm{s}} = 4.5$ (4.3), $C_{\mathrm{k}} = 36.4$ (30.2) and probability dry 44% (48%), where inside parentheses are the values of the observations, which are relatively well preserved. For illustration purposes, in Fig. 4 we plot a 3000 days window of the observed vs. the simulated precipitation. Note that here, the explicit preservation up to the fourth moment is adequate, since preservation of additional moments slightly improve the distribution simulation (specifically, the R^2 coefficient is estimated for preservation of the 1st, 2nd, 3rd, 4th, 5th, and 6th moment as 0.953, 0.985, 0.985, 0.9861, 0.9863 and 0.9864, respectively).

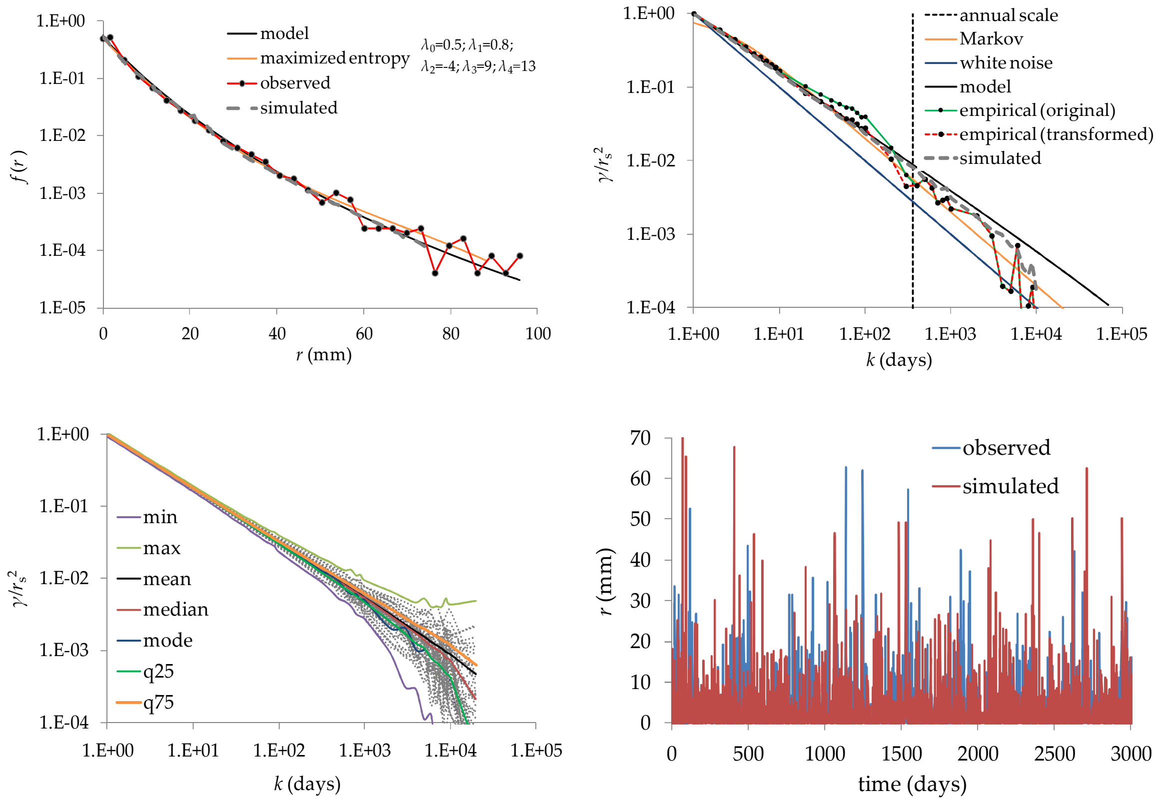

Figure 4: [upper left] Empirical (original and transformed to approximately remove seasonality), modelled and simulated marginal distributions (corresponding weights for the ME distribution: 73%, 15%, 7% and 5%); [upper right] climacograms for the standardized precipitation process; [lower left] the mode and several other essential statistical measures of the standardized climacograms estimated from 10^3 synthetic time series (in the figure we depict only 50 empirical climacograms); [lower right] a 3000 days window of the observed precipitation record along with a simulated one.

4.3. Application to hourly surface wind; a GHK process with a three-parameter distribution

For the hourly wind process we adopt the GHK process (Eqn. 8) for the dependence structure. For the distribution function we apply a special case of the PBF distribution which approximates the Weibull distribution for small hourly velocities and the Pareto distribution for larger ones (e.g., Aksoy et al., 2004; Lo Brano et al., 2011). The dependence structure, marginal distribution and standardization scheme of wind are based on the preliminary analysis from thousands of stations around the globe, performed by Dimitriadis and Koutsoyiannis (2016). A more thorough analysis justifying the above choices for the wind process can be seen in Koutsoyiannis et al. (2018) and references therein. The three-parameter GHK process and selected PBF marginal probability function can be written as:

$$\gamma(k) = \frac{\lambda}{(1 + k/q)^{2-2H}} \tag{18}$$

$$F(v) = 1 - \left(1 + \left(\frac{v}{\alpha v_s}\right)^b\right)^{-c/b} \tag{19}$$

where $v > 0$ is the wind process; $k = \kappa D$ is the continuous time scale with D = 1 h the sampling time interval and κ the discrete time scale; q is the scale parameter of the process; λ is the true variance of the instantaneous (continuous-time) process; H is the Hurst parameter; v_s is the standard deviation of the discretized process that should equal the expected value of the square root of the climacogram for scale κ = 1, i.e. $v_s = \sqrt{\gamma(D)} = (1 + D/q)^{H-1}\sqrt{\lambda}$; in addition, α is the scale parameter and b and c are the shape parameters of the marginal distribution, all dimensionless. Interestingly, here the survival function and the dependence structure have identical expressions (Koutsoyiannis, et al., 2018). Note that by assuming stationarity and ergodicity, we are able to standardize the wind process (Fig. 5), in order to homogenize all time series recorded at different locations, altitude and climatic conditions (this should not be confused with normalization through a non-linear transformation).

We choose to apply the above stochastic model to the longest nine hourly wind time series of different lengths located in Greece (Table 1). The expression for the bias of the classical estimator of the climacogram is derived in Tyralis and Koutsoyiannis (2011) for an HK process and generalized for all processes in Koutsoyiannis (2011). Here, we use the general expression and, since the time series have different lengths n, we apply an estimator of the climacogram adjusted for n:

$$\underline{\hat{\gamma}}(\kappa D) = \frac{1}{\lfloor n/\kappa \rfloor} \sum_{i=1}^{\lfloor n/\kappa \rfloor} \left(\underline{\overline{x}_l} - \underline{\overline{x}}\right)^2 + \gamma(\lfloor n/\kappa \rfloor \kappa D) \tag{20}$$

where $\underline{\hat{\gamma}}(\kappa D)$ is an unbiased estimator of the climacogram $\gamma(kD)$, since $\mathrm{E}\left[\underline{\hat{\gamma}}(\kappa D)\right] = \gamma(\kappa D)$.

Table 1: General information of the meteorological stations and statistical characteristics of the hourly wind time series (downloaded from ftp.ncdc.noaa.gov).

hourly wind station	longitude (deg)	latitude (deg)	elevation above sea level (m)	no. of years	mean (m/s)	st. deviation (m/s)	missing values (%)	zero values (%)
Heraκleio	25.183	35.333	39	39	4.583	2.918	8.8	6.3
N. Aghialos	22.8	39.217	15	17	3.258	2.331	28	19
Karpathos	35.417	27.15	20	17	7.506	4.074	30.4	3.9
Santorini	36.4	25.483	38	24	5.701	3.229	29.5	7.5
Kos	36.8	27.083	125	33	4.805	2.7	15	7
El. Venizelos	37.93	23.93	96	11	3.954	2.995	0.6	1.9
Limnos	39.917	25.233	5	38	4.458	3.546	23	17.5
Paros	37.02	25.13	36	11	5.567	3.265	46.8	6.5
Meganissi	38.95	20.767	4	40	3.571	2.746	36.3	19.4

The parameters related to the dependence structure via the climacogram are estimated as: $\lambda = 1.3$, $q = 5$ h and $H = 0.75$, whereas for the marginal distribution as: $a = 6$, $b = 1.9$ and $c = 14.8$, corresponding to $\mu = 1.9$, $\sigma = 1.1$ ($\approx \sqrt{\lambda}$), $C_s = 1.2$ and $C_k = 4.8$ (all estimations are based on the fitting norms in Eqns. 14 and 16). Again, the explicit preservation up to the fourth moment is adequate, since preservation of additional moments slightly improve the distribution simulation (specifically, the R^2 coefficient is estimated for preservation of the 1st, 2nd, 3rd, 4th, 5th, and 6th moment as 0.936, 0.949, 0.977, 0.983, 0.984, and 0.984, respectively). To emulate the observed wind time series one could set to zero any values of the synthetic time series that are below the corresponding recording threshold of a typical anemometer, which is around 0.5 m/s depending on the type of the anemometer (e.g. Conradsen et al., 1984). For illustration purposes, in Fig. 5 we plot a 1000-day window of the observed vs. the simulated wind speed at Kos Island. The empirical and modelled probability of standardized wind speed less than or equal to 1 are both around 50%. Note that σ and λ should approximate unity but they are slightly larger due to the double cyclostationary effects of the daily and seasonal periodicities of the wind process (Deligiannis et al., 2016). These effects cause the small increase of climacogram around daily and annual scales (Fig. 5) but here, for simplicity, we apply a stationary rather than a cyclostationary model through the non-linear transformation of the probability function of Deligiannis et al. (2016). Again, due to the weak periodicities of the examined process the double cyclostationarity can be generated through the inverse transformation.

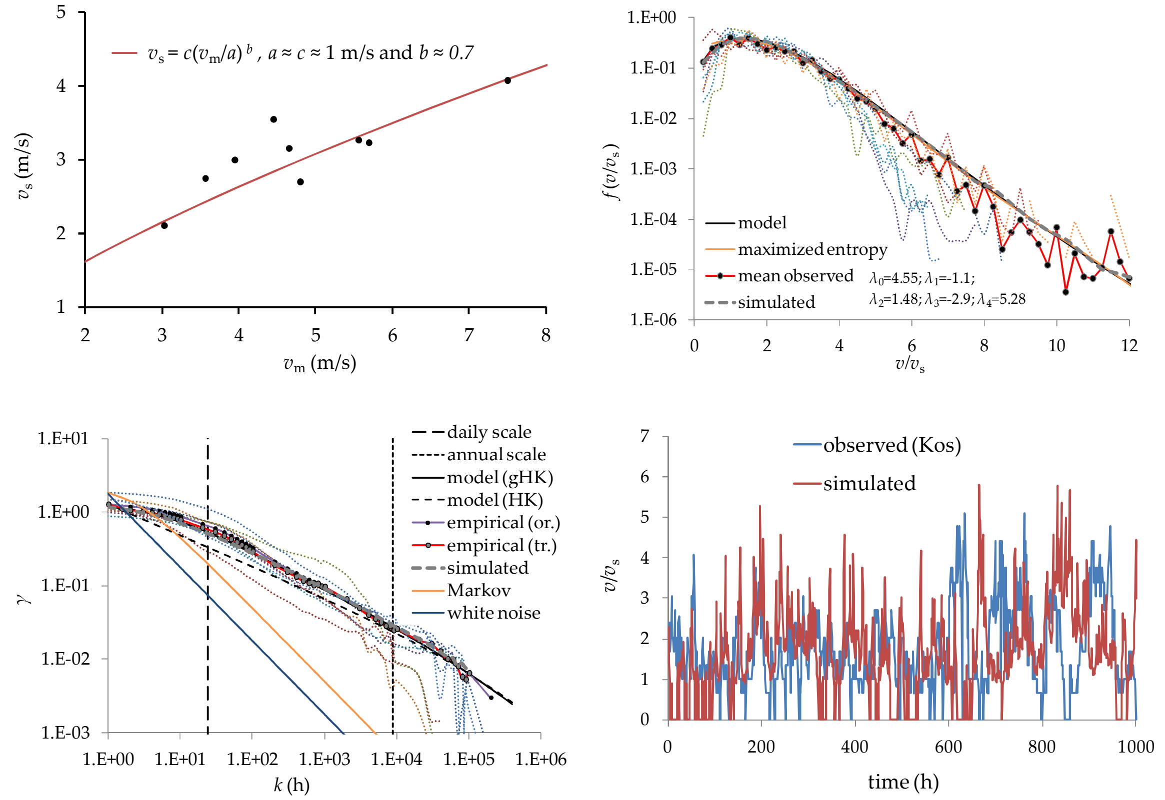

Figure 5: Empirical mean (v_m) vs. standard deviation of the nine time series along with the fitted model [upper left]; the empirical (original and transformed to approximately remove the double periodicity), model and simulated marginal distributions (corresponding weights for the ME distribution: 43%, 32%, 16% and 9%) [upper right] and climacograms [lower left] for the standardized wind process; a 1000-day window of the observed standardized wind process in Kos island along with a standardized simulated one [lower right].

4.4. Application to turbulence; an HHK process with an unknown distribution

As already mentioned, high-order moments cannot be reliably estimated from typically short time series of geophysical processes. However, in laboratory experiments with high sampling rates, very large time series of observations can be formed, which allow direct estimation of high order moments from data. Here, we use a grid-turbulence massive database provided by the Johns Hopkins University (www.me.jhu.edu/meneveau/datasets/datamap.html). This dataset includes 40 time series, each with n = 36×10^6 data points of longitudinal wind velocity along the flow direction, all measured by X-wire probes placed downstream of the grid and with a sampling time interval of 25 μs (Kang et al., 2003). Due to the laboratory nature of the experiment we may apply the Taylor's hypothesis of frozen turbulence (Taylor, 1938) and shift from the spatial to the temporal domain (Castro et al., 2011). We then use a standardization scheme illustrated in Fig. 6 to homogenize all series (Dimitriadis et al., 2016a) and, by setting the empirical mean to zero, we calculate the standardized empirical variance as $\mathrm{E}[\underline{\hat{\gamma}}(D)] \approx 1$. By the standardization, we are able to form a sample of 40 × 36 ×10^6 = 1.44 ×10^9 values for the estimation of the marginal characteristics of the process and an ensemble of 40 series, each with 36 ×10^6 values for the estimation of the dependence structure characteristics.

It can be observed that the time series are not Gaussian but rather nearly-Gaussian as shown in Fig. 7. This is also verified by the skewness and kurtosis estimates of 0.2 and 3.1, respectively. If those values were estimated from a small sample, for example n = 100, then the probability density function of the process would be regarded Gaussian and the divergence from normality would be attributed to statistical error, since for n = 100 the uncertainty measured through the standard deviation of the skewness and kurtosis, is as high as 30% and 50%, respectively (Fig. 2). However, for $n \approx 1.5\times10^9$ the uncertainty of the mean will drop below 1% for H = 0.8 and therefore, based on extrapolation of curves in Fig. 2, it is seen that the uncertainty of skewness and kurtosis will be low too. Moreover, there are theoretical arguments justifying the divergence of fully developed turbulent processes from normality (Wilczek et al., 2011).

In contrast to the earlier application, where the value of the fourth moment was inferred from the fitted theoretical distribution, here we directly estimate it from the data since the estimation error is very low due to the huge number of data points as well as due to the high quality of the measurements.

For the stochastic structure, we apply a stochastic model (modified HHK process) and fit it by incorporating both discretization and bias effects (Fig. 8). This model combines both fractal and HK dynamics using four parameters and it attributes the grid-turbulent process to an HHK process (Eqn. 8) and the independent effect of the boundaries of the experiment, which cause a drop of variance at intermediate scales, to a Markov process:

$$\gamma(k) = \frac{\lambda}{2\left(1+(k/q)^{2M}\right)^{\frac{1-H}{M}}} + \frac{\lambda\left(k/q + \mathrm{e}^{-k/q} - 1\right)}{(k/q)^2} \tag{21}$$

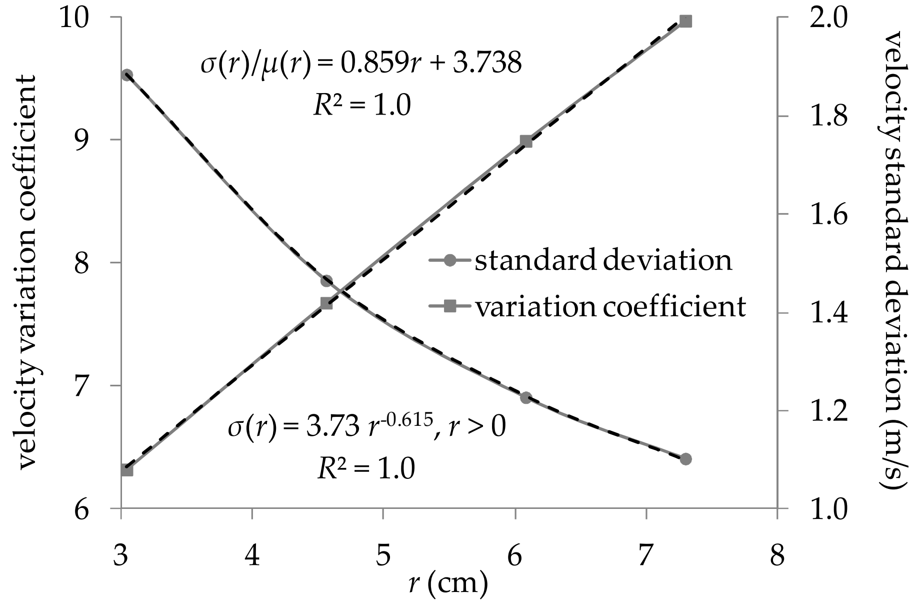

Figure 6: Standardization scheme for grid-turbulence data, where μ and σ are the mean and standard deviation, r is the distance from the grid, with the first 16 time series corresponding to transverse points abstaining r = 20L from the source, the second 4 to r = 30L, the third 4 to 40L and the last 16 to 48L, with L = 0.152 m the size of the grid.

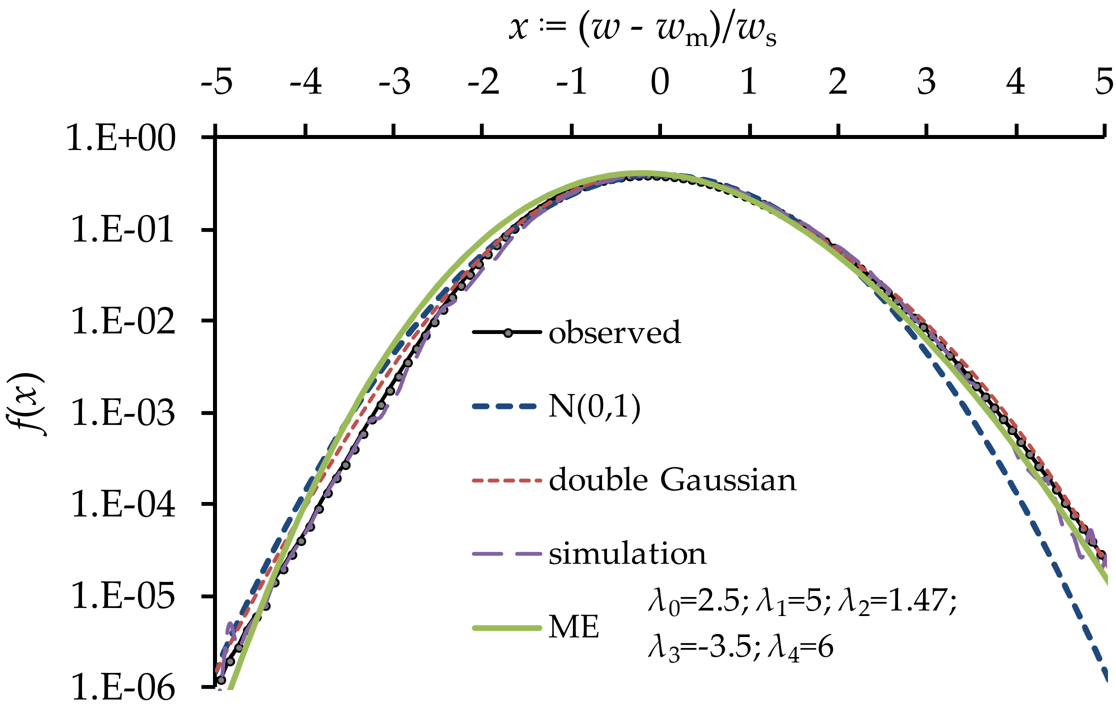

Figure 7: Empirical probability density function of the overall standardized time series (observed) along with that from a single synthetic time series produced by the SMA scheme to preserve the first four moments (simulation); for comparison the theoretical distributions N(0,1), sum of two Gaussian distributions (double Gaussian), and ME constrained on the four moments (corresponding weights for the ME distribution are estimated as 15%, 51%, 21% and 13%.

Again here, the explicit preservation up to the fourth moment is adequate, since preservation of additional moments slightly improve the distribution simulation (specifically, the R^2 coefficient is estimated for preservation of the 1st, 2nd, 3rd, 4th, 5th, and 6th moment as 0.0372, 0.990, 0.991, 0.998, 0.999, and 0.999, respectively).

For the estimation of the climacogram we use the estimator of Eqn. (17) and we apply the methodology by fitting the expected model to the mean climacogram calculated from the 36 time series of identical length. However, to improve the fitting of the model, we include in the analysis the climacogram-based structure function (abbreviated CBF) and the climacogram-based spectrum (abbreviated CBS), as introduced in Koutsoyiannis (2016). The climacogram is more representative of the large and intermediate scales, the CBF of the small and intermediate scales and the CBS of small and large scales and thus, by combining all three of them we can achieve a better fitting of the model (Dimitriadis et al., 2016a). The CBF and CBS are defined through the climacogram respectively, as:

$$\xi(k) := \gamma(0) - \gamma(k) \tag{22}$$

$$\psi(w) := \frac{2}{w\gamma(0)}\gamma(1/w)\xi(1/w) = \frac{2\gamma(1/w)}{w}\left(1 - \frac{\gamma(1/w)}{\gamma(0)}\right) \tag{23}$$

where $\gamma(0) = \lambda$ for the specified mode; and $w = 1/k$ is the frequency for a continuous-time process (in inverse time units).

The model parameters are estimated as: $\lambda = 1$, $M = 1/3$, $H = 5/6$ and $q = 14$ ms, through the fitting norm of Eqn. 16. Here a large number of parameters could be justified due the large data size but the above model is quite parsimonious (it has two parameters less than that used by Dimitriadis et al., 2016a, for modelling the same process). Also, the applied extended HHK model is theoretically justified through the maximization of entropy (Koutsoyiannis, 2011) and therefore, each parameter has a physically-based interpretation. Moreover, we observe from Fig. 8 that this model comes also in agreement with the work on the turbulent power spectrum by von Karman (1948) for the large scale range, by Kolmogorov (1941a-c; K41 model) for the intermediate range and by Kraichnan (1959) for the dissipation range (cf. Pope 2000, pp. 232-233), while here we also simulate the Hurst-Kolmogorov behaviour that clearly appears in the very small frequencies (very large scales) of the power spectrum and in the other stochastic tools in Fig. 8. Additionally, certain aspects exhibited in the power spectrum such as the bottleneck effect (Kang et al., 2003) and the spike at large frequencies (which is often ignored and attributed to instrumental noise; Cerutti and Meneveau, 2000) are also well represented. Finally, the preservation of kurtosis of the velocity increments enables to even simulate the effect that the intermittent behaviour of the process has on the marginal probability distribution, first discovered in turbulence by Batchelor and Townsend (1949).

It is interesting to further investigate the latter issue through the behaviour of a generalized structure function $V_p(h) := \mathrm{E}[|\underline{x}_i - \underline{x}_{i+h}|^p]$ and in particular the power-law behaviour for the intermediate range of lags, i.e. $V_p(h) \approx h^{\zeta_p}$. Such behaviours have been attributed to intermittency (Frisch, 2006, sect. 8.3) which initiated the need for exploring models different from the K41 such as the multifractal ones (Frisch, 2006, sect. 8.5 to 8.9). As shown in Fig. 9 to 12, the drop of skewness (Fig. 9) and the drop of kurtosis (Fig. 10) of the velocity increments for a wide range of lag (h) or the regular velocity vs scale (high order climacogram; Fig. 11) are impressively well preserved by the proposed model. It is important to notice in Figs. 9 and 10 that, if preservation up to third (rather than fourth) marginal moment was made, then the lagged skewness and kurtosis would not be preserved adequately. In addition, the increase of the exponent ζ_p is equally well preserved by the proposed model for a wide range of the p exponent (Fig. 12). This is achieved with no particular effort or provision (e.g., without using extra assumptions, parameters or models) but merely by simultaneously simulating the first four moments (with focus on the coefficient of kurtosis) and the stochastic structure of the process. To further highlight this finding, we illustrate in Fig. 12 that the HHK model alone cannot simulate the observed behaviour of the high order structure function but rather approaches the structure function as simulated by the K41 self-similarity model and reproduced by Frisch (2006, Fig. 8.8). Similar results are obtained in case a Markov dependence structure is adopted but by simultaneously preserving the empirical non-Gaussian marginal distribution. Interestingly, if both the proposed dependence structure and marginal distribution are combined, then the observed behaviour of the high order structure function is approximated and, as a consequence, the intermittent behaviour of turbulence. For comparison, in Fig. 12 we plot the She-Leveque model (1994) that behaves also exceptionally well and originates from the alternative assumption of independent identically distributed log-Poisson multiplicative factors (Frisch, 2006, sect. 8.6.4, 8.6.5).

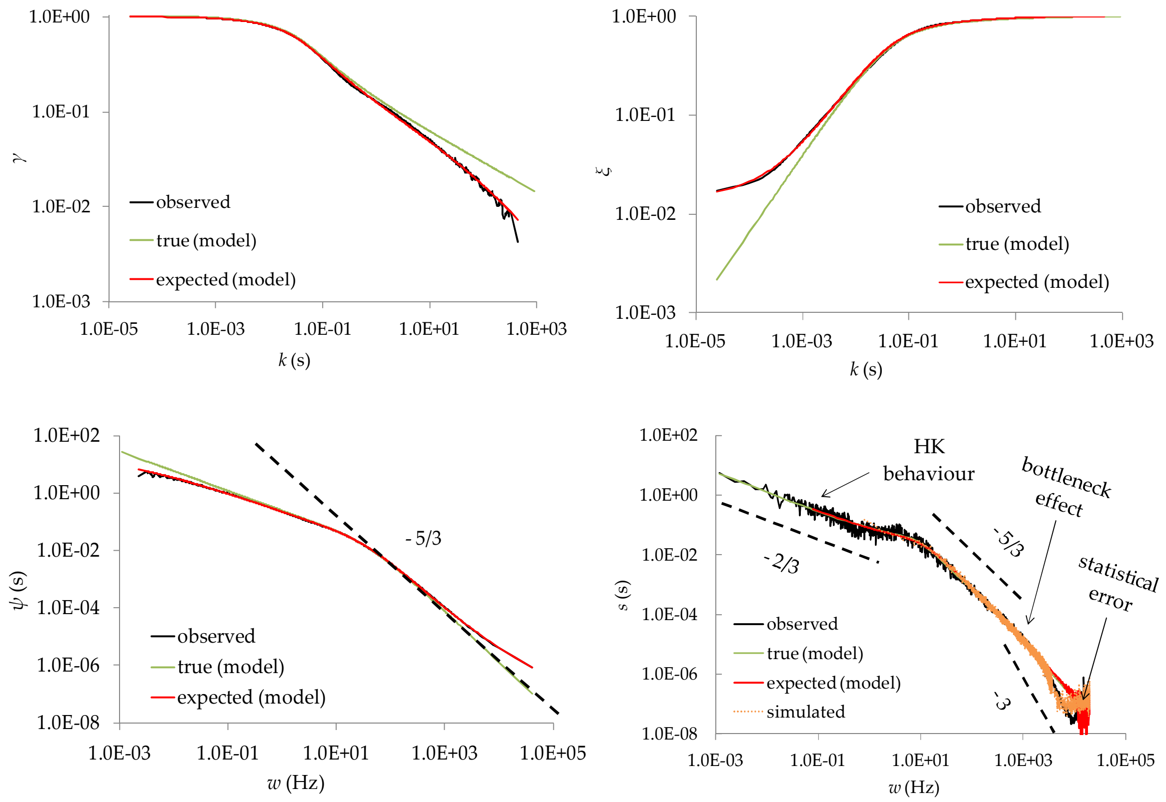

Figure 8: The empirical, true and expected values of the climacogram [upper left], CBF [upper right], CBS [lower left] and power spectrum [lower right] along with some important logarithmic slopes; and with a correlation coefficient (as a rough indicator of the goodness of fit) between the modelled and simulated values estimated as 99.99%, 99.99%, 99.56%, and 94.08%, for each metric, respectively.

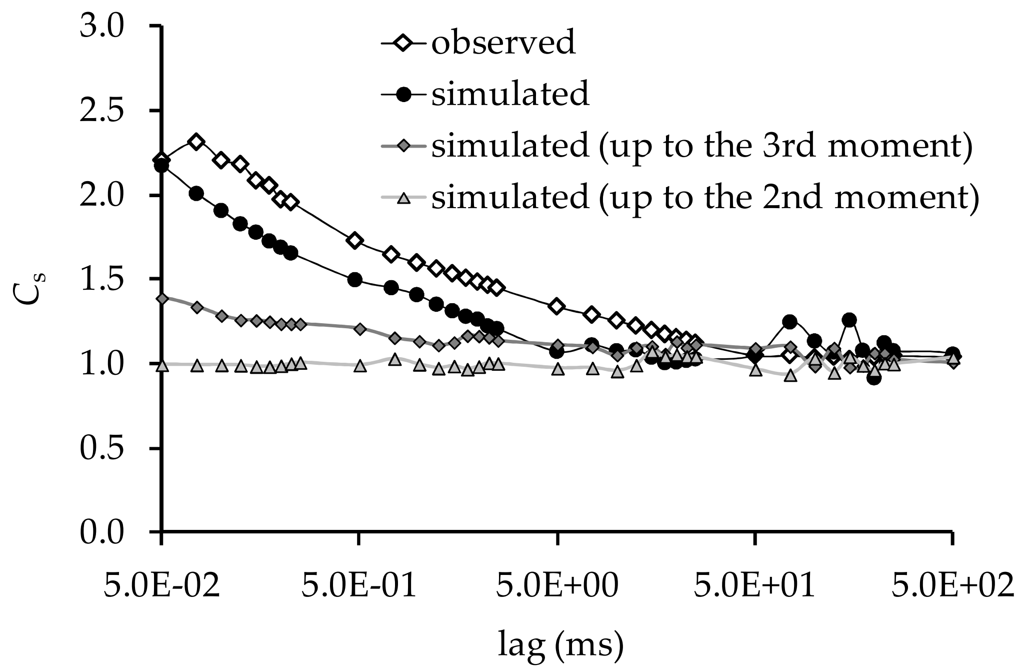

Figure 9: Empirical and simulated skewness coefficient of the first order structure function vs. lag.

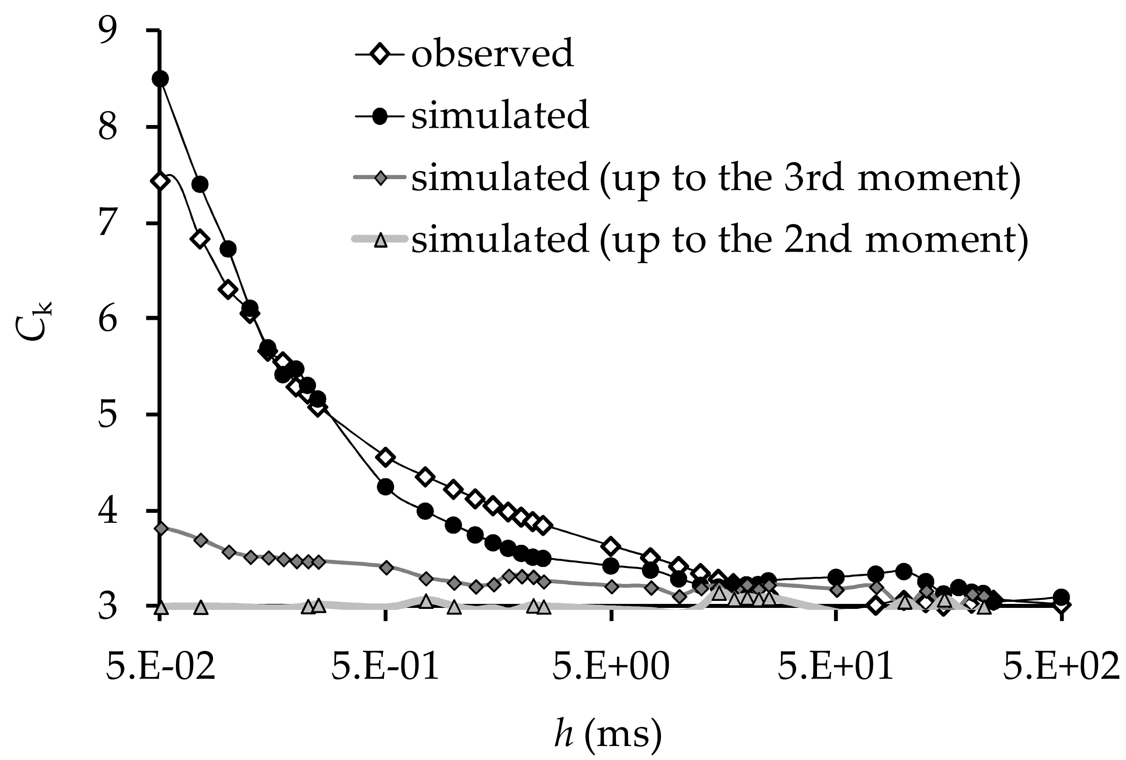

Figure 10: Empirical and simulated kurtosis coefficient of the velocity increments vs. lag.

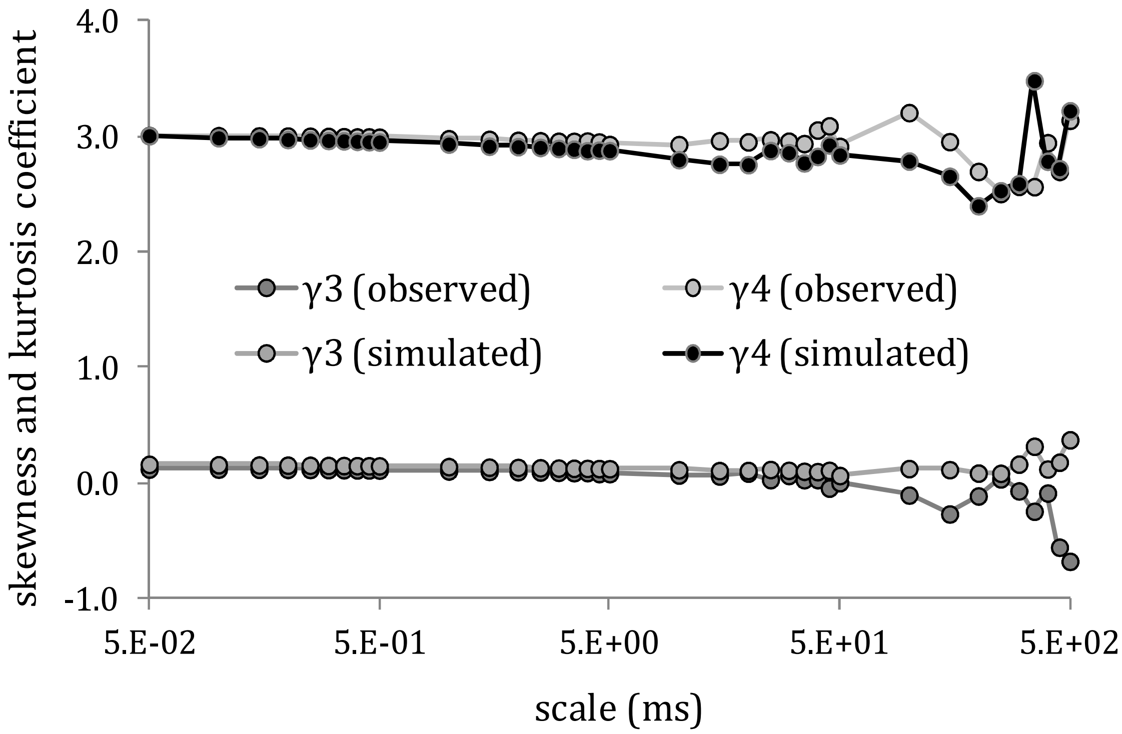

Figure 11: High-order climacograms: coefficients of skewness (γ_3) and kurtosis (γ_4) vs. scale (empirical vs. simulated).

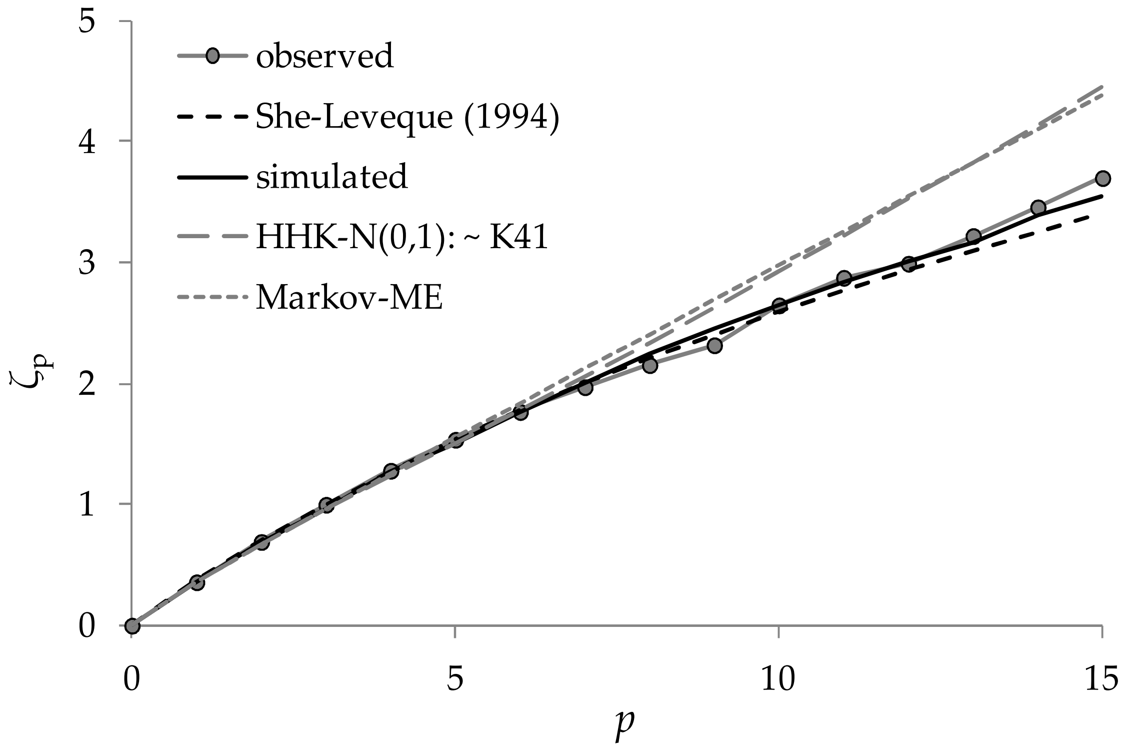

Figure 12: Empirical and simulated (through the She-Leveque model, the SMA scheme, a Gaussian HHK model and a Markov model with the ME distribution) structure function for various orders of the velocity increments vs. lag.

5. Conclusions

We present an extension of the SMA framework to include representation of high-order moments and in particular the simulation of the kurtosis coefficient of the process. In this way, the probability density function of any process is approximated by its first four central moments, thus, according to the maximum entropy approach, yielding a probability density function that is an exponentiated fourth order polynomial. In the generation phase, the approximation is performed using convenient four-parameter distributions (a standardized Kumaraswamy distribution for thin-tailed distributions, or a standardized Normal-Inverse-Gaussian distribution for heavy-tailed ones). Application to non-Gaussian white noise with various customary distributions shows that the approximation achieved is very satisfactory.

The presented scheme is very useful for stochastic generation as well as Monte-Carlo experiments (for sensitivity analyses, for the derivation of confidence intervals etc.), especially when numerous measurements exist (e.g. in laboratory experiments), while a theoretical distribution cannot be easily identified but several statistical moments can be estimated. The limitation of this methodology is that the marginal distribution is approximated to a desired degree, rather than precisely preserved (particularly in non-divisible distributions). This limitation may create difficulties in variables with upper or lower bounds, since these can be only treated in an ad-hoc manner. However, this limitation rarely concern practical applications to geophysical processes.

In this work, we apply the methodology to a long daily precipitation record, to various record lengths of hourly surface wind time series, and to a grid-turbulence massive database. In all applications the extended SMA scheme performs exceptionally well, additionally preserving other important statistical characteristics of the processes such as the intermittent behaviour. Particular emphasis has been given on turbulence, in an attempt to show that several aspects of turbulence regarded as puzzles can be easily reproduced by a simple unpuzzling model without a particular effort.

Additional contributions of this paper are the estimators of the dependence structure of a process, accounting for the statistical bias, in the case of the analysis of a single time series (as in the application of precipitation) and of several time series of the same process with different lengths (as in the application of wind) and identical lengths (as in the application of grid-turbulence). Also, we note the very good fit of special cases of the PBF distribution to the marginal distribution of the precipitation and wind processes.

Acknowledgements

The Authors are grateful to the Editor George Christakos, the anonymous Associate Editor and three Reviewers as well as to T. Iliopoulou, for their efforts and useful comments and suggestions that helped us improve the paper.

Funding information

No funds have been available for this work.

Code availability

The SMA model used in this paper as well as a fast code for estimation of the climacogram for long time series, both implemented in Matlab can be downloaded from www.itia.ntua.gr/1656/. In this link, we also provide scripts for various stochastic models, such as the HK, GHK and HHK as well as the applied models in section 4 for the precipitation, wind speed and grid-turbulence processes.

References

Aksoy, H., Toprak, Z.F., Aytek, A., Ünal, N.E., Stochastic generation of hourly mean wind speed data, *Renew. Energy*, 29, 2111–2131, 2004.

Balestrino, A., A. Caiti and E. Crisostomi, Efficient numerical approximation of maximum entropy estimates, *International Journal of Control*, 79(9):1145-1155, 2006.

Barndorff-Nielsen, O., Hyperbolic Distributions and Distributions on Hyperbolae, *Scandinavian Journal of Statistics*, 1978.

Batchelor, G.K., and Townsend, A.A., The nature of turbulent motion at large wave-numbers, *Proc. R. Soc. Lond.*, A 199, 238-255, 1949.

Castro, J.J., Carsteanu, A.A., and Fuentes, J.D., On the phenomenology underlying Taylor's hypothesis in atmospheric turbulence, *Revista Mexicana de Fisica*, 57(1) 60–64, 2011.

Cerutti, S. and Meneveau, C., Statistics of filtered velocity in grid and wake turbulence. *Physics of Fluids*, 12(1): 143-1 165, 2000.

Chhikara, R., and Folks, L., The inverse Gaussian distribution: theory, methodology and applications. *New York: Marcel Dekker*, 1989.

Conradsen, K., Nielsen, L.B., and Prahm, L.P., Review of Weibull statistics for estimation of wind speed distributions, *J. Clim. Appl. Meteorol.*, 23, 1173–1183, 1984.

Cordeiro, G.M., and de Castro, M., A new family of generalized distributions, *J. Stat. Comput. Simul.*, 81, 883–898, 2011.

Gugar, V.G. and R.J. Kavanagh, Generation of random signals with specified probability density functions and power density spectra, *IEEE-AC*, 13 :716–719, 1968.

Deligiannis, H., Dimitriadis, P., Daskalou, O., Dimakos, Y., and Koutsoyiannis, D., Global investigation of double periodicity of hourly wind speed for stochastic simulation; application in Greece, *Energy Procedia*, 97, 278–285, doi:10.1016/j.egypro.2016.10.001, 2016.

Dimitriadis, P., Hurst-Kolmogorov dynamics in hydrometeorological processes and in the microscale of turbulence, PhD thesis, 167 pages, National Technical University of Athens, Athens, 2017.

Dimitriadis, P., and D. Koutsoyiannis, Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes, *Stochastic Environmental Research & Risk Assessment*, 29 (6), 1649–1669, 2015a.

Dimitriadis, P., and D. Koutsoyiannis, Application of stochastic methods to double cyclostationary processes for hourly wind speed simulation, *Energy Procedia*, 76, 406–411, doi:10.1016/j.egypro.2015.07.851, 2015b.

Dimitriadis, P. and D. Koutsoyiannis, A parsimonious stochastic model for wind variability, *30 Years of Nonlinear Dynamics in Geosciences*, Rhodes, Greece, 2016.

Dimitriadis, P., D. Koutsoyiannis, and P. Papanicolaou, Stochastic similarities between the microscale of turbulence and hydrometeorological processes, Hydrological Sciences Journal, 61 (9), 1623–1640, doi:10.1080/02626667.2015.1085988, 2016a.

Dimitriadis, P., D. Koutsoyiannis, and K. Tzouka, Predictability in dice motion: how does it differ from hydrometeorological processes?, *Hydrological Sciences Journal*, 61 (9), 1611–1622, 2016b.

Dimitriadis, P., N., Gournary and D. Koutsoyiannis, Markov vs. Hurst-Kolmogorov behaviour identification in hydroclimatic processes, *European Geosciences Union General Assembly 2016, Geophysical Research Abstracts, Vol. 18*, Vienna, EGU2016-14577-4, European Geosciences Union, 2016c.

Dimitriadis, P., T., Iliopoulou, H., Tyralis, and D., Koutsoyiannis, Identifying the dependence structure of a process through pooled time series analysis, *IAHS 2017 Scientific Assembly 10 – 14 JULY 2017, Port Elizabeth, South Africa, Water and Development: scientific challenges in addressing societal issues, W17: Stochastic hydrology: simulation and disaggregation models*, 2017.

Efstratiadis, A., Dialynas, Y., Kozanis, S., and Koutsoyiannis D., A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence, Environmental Modelling and Software, 62, 139–152, doi:10.1016/j.envsoft.2014.08.017, 2014.

Fernandez, B. and Salas, J.D., Periodic Gamma Autoregressive Processes for Operational Hydrology, *Water Resour. Res.*, 22(10):1385-1396, 1986.

Frechet M., Sur les tableaux de correlation dont les marges son donnees, *Ann Univ Lyon*, Sect A 9:53–77, 1951.

Frisch, U., Turbulence: The Legacy of A. N. Kolmogorov, *Cambridge University Press*, Cambridge, 2006.Halliwell, Leigh J., Classifying the Tails of Loss Distributions, *CAS E-Forum*, Vol. 2, 2013.

Gneiting T, Power-law correlations, related models for long-range dependence and their simulation, *J Appl. Prob.*, 37(04):1104–1109, 2000.

Gneiting T., and Schlather M., Stochastic models that separate fractal dimension and the hurst effect, *SIAM Rev* 46(2): 269–282, 2004.

Gneiting T., Evcıkova, H., and Percival D.B., Estimators of fractal dimension: assessing the roughness of time series and spatial data, *Stat Sci* 27(2): 247–277, 2012.

Halliwell L.J., Classifying the tails of loss distributions, *CAS EForum*, vol 2, 2013.

Hoeffding, W., *Scale-invariant correlation theory*, in N. I. Fisher and P. K. Sen (Eds.), The Collected Works of Wassily Hoeffding, pp. 57–107, New York: Springer-Verlag, 1940.

Ibragimov R., and Lentzas G., Copulas and long memory, *Probab. Surveys*, 14:289–327. https://doi.org/10.1214/14-PS233, 2017.

Iliopoulou, T., S.M. Papalexiou, Y. Markonis, and D. Koutsoyiannis, Revisiting long-range dependence in annual precipitation, *Journal of Hydrology*, 2016.

Jaynes, E.T., Information Theory and Statistical Mechanics, *Phys. Rev.,* 106**,** 620, 1957.

Kang, H.S., Chester, S. and Meneveau, C., Decaying turbulence in an active-grid-generated flow and comparisons with large-eddy simulation, *Journal of Fluid Mechanics,* 480, 129-160, 2003.

Khan, M.S., King. R., and Hudson, I.L., *STATISTICS IN TRANSITION new series*, Vol. 17, No. 2, pp. 1–28, Transmuted Kumaraswamy Distribution, 2016.

Klugman, S.A., Panjer, H.H., and Willmot, G.E., Loss Models: From Data to Decisions, *John Wiley & Sons*, New York, 1998.

Kolmogorov, A. N., Uber die analytischen Methoden in der Wahrscheinlichkcitsrechnung, Math. Ann. 104, 415-458. (English translation: On analytical methods in probability theory, In: Kolmogorov, A.N., 1992. Selected Works of A. N. Kolmogorov - Volume 2, *Probability Theory and Mathematical Statistics*, A. N. Shiryayev, ed., Kluwer, Dordrecht, The Netherlands, pp. 62-108), 1931.

Kolmogorov, A. N., *Grundbegrijfe der Wahrscheinlichkeitsrechnung*, Ergebnisseder Math. (2), Berlin. (2nd English Edition: Foundations of the Theory of Probability, 84 pp. Chelsea Publishing Company, New York, 1956), 1933.

Kolmogorov, A.N., The local structure of turbulence in incompressible viscous fluid for very large Reynolds number, *Doklady Akademii Nauk SSSR*, 30, 299-303, 1941a.

Kolmogorov, A.N., On the decay of isotropic turbulence in an incompressible viscous flow, *Doklady Akademii Nauk SSSR*, 31, 538-540, 1941b.

Kolmogorov, A.N., Dissipation energy in locally isotropic turbulence, *Doklady Akademii Nauk SSSR*, 32, 16-18, 1941c.

Koudouris, G., P. Dimitriadis, T. Iliopoulou, N. Mamassis, and D. Koutsoyiannis, Investigation on the stochastic nature of the solar radiation process, *Energy Procedia*, 125, 398–404, 2017.

Koutsoyiannis, D., A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research*, 36 (6), 1519–1533, 2000.

Koutsoyiannis, D., The Hurst phenomenon and fractional Gaussian noise made easy, Hydrological Sciences Journal, 47 (4), 573–595, 2002.

Koutsoyiannis, D., Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, 48 (1), 3–24, 2003.

Koutsoyiannis, D., Statistics of extremes and estimation of extreme rainfall: I. Theoretical Investigation, *Hydrol. Sci. J.*, 49, 2004a.

Koutsoyiannis, D., Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records, *Hydrol. Sci. J.*, 49, 2004b.

Koutsoyiannis, D., Uncertainty, entropy, scaling and hydrological stochastics, 1, Marginal distributional properties of hydrological processes and state scaling, *Hydrological Sciences Journal*, 50 (3), 381–404, 2005.

Koutsoyiannis, D., HESS opinions "A random walk on water", *Hydrology and Earth System Sciences*, 14, 585–601, 2010.

Koutsoyiannis, D., Hurst-Kolmogorov dynamics as a result of extremal entropy production, *Physica A: Statistical Mechanics and its Applications*, 390 (8), 1424–1432, 2011.

Koutsoyiannis, D., Entropy: from thermodynamics to hydrology, *Entropy*, 16 (3), 1287–1314, doi:10.3390/e16031287, 2014.

Koutsoyiannis, D., Generic and parsimonious stochastic modelling for hydrology and beyond, *Hydrological Sciences Journal*, 61 (2), 225–244, 2016.

Koutsoyiannis D., Entropy production in stochastics, *Entropy*, 19(11): 581, 2017.

Koutsoyiannis D., and Manetas A., Simple disaggregation by accurate adjusting procedures, *Water Resour. Res.*, 32(7): 2105–2117, https://doi.org/10.1029/96WR00488, 1996.

Koutsoyiannis, D., and A. Montanari, Negligent killing of scientific concepts: the stationarity case, *Hydrological Sciences Journal*, 60 (7-8), 1174–1183, 2015.

Koutsoyiannis, D., Onof, C., and Wheater, H.S., Multivariate rainfall disaggregation at a fine timescale, *Water Resources Research*, 39 (7), 1173, doi:10.1029/2002WR001600, 2003.

Koutsoyiannis, D., Yao, H., and Georgakakos, A., Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods, *Hydrological Sciences Journal*, 53 (1), 142–164, 2008.

Koutsoyiannis, D., Dimitriadis, P., Lombardo, F., and Stevens, S., From fractals to stochastics: Seeking theoretical consistency in analysis of geophysical data, *Advances in Nonlinear Geosciences*, Tsonis, A. ed.; Springer, New York, USA, 2018; doi: 10.1007/978-3-319-58895-7, 2018.

Kraichnan, R.H., Stochastic modeling of isotropic turbulence, *Springer*, 1991.

Krajewski, W. F., Kruger, A., and Nespor, V., Experimental and numerical studies of small-scale rainfall measurements and variability, *Water Sci. Technol.*, 37, 131–138, 1998.

Kumaraswamy, P., A generalized probability density function for double-bounded random processes, *Journal of Hydrology*, 46 (1-2): 79–88, 1980.

Langousis, A., and Koutsoyiannis, D., A stochastic methodology for generation of seasonal time series reproducing overyear scaling behaviour, *Journal of Hydrology*, 322, 138–154, 2006.

Lavergnat, J., On the generation of colored non-Gaussian time sequences, *hal* (01399446), 2016.

Lebrun, R. and A. Dutfoy, An innovating analysis of the Nataf transformation from the copula viewpoint, *Probabilistic Engineering Mechanics*, 24(3): 312-320, 2009.

Lo Brano V., Orioli A., Ciulla G., and Culotta S., Quality of wind speed fitting distributions for the urban area of Palermo, Italy, *Renew Energy*, 36(3):1026–1039, 2011.

Lombardo, F., Volpi, E., and Koutsoyiannis, D.: Rainfall downscaling in time: theoretical and empirical comparison between multifractal and Hurst-Kolmogorov discrete random cascades, *Hydrolog. Sci. J.*, 57, 1052–1066, 2012.

Lombardo, F., E. Volpi, D. Koutsoyiannis, and S.M. Papalexiou, Just two moments! A cautionary note against use of high-order moments in multifractal models in hydrology, *Hydrology and Earth System Sciences*, 18, 243–255, 2014.

Lombardo, F., E. Volpi, D. Koutsoyiannis, and F. Serinaldi , A theoretically consistent stochastic cascade for temporal disaggregation of intermittent rainfall, *Water Resources Research*, doi:10.1002/2017WR020529, 2017.

Markonis, Y., and D. Koutsoyiannis, Climatic variability over time scales spanning nine orders of magnitude: Connecting Milankovitch cycles with Hurst–Kolmogorov dynamics, Surveys in Geophysics, 34 (2), 181–207, 2013.

Moschos, E., G. Manou, P. Dimitriadis, V. Afendoulis, D. Koutsoyiannis, and V. Tsoukala, Harnessing wind and wave resources for a Hybrid Renewable Energy System in remote islands: a combined stochastic and deterministic approach, *Energy Procedia*, 125, 415–424, doi:10.1016/j.egypro.2017.08.084, 2017.

Nataf A., Statistique mathematique-determination des distributions de probabilites dont les marges sont donnees, *C R Acad Sci*, Paris, 255:42–43, 1962.

Nelsen, R.B., *An Introduction to Copulas*, Springer Series in Statistics, second edition, 2006.

Nespor, V., and Sevruk, B., Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation, *J. Atmospheric and Oceanic Technol.*, 16, 450–464, 1999.

O'Connell, P.E., Koutsoyiannis, D., Lins, H.F., Markonis, Y., Montanari, A., Cohn, T., The scientific legacy of Harold Edwin Hurst (1880–1978), *Hydrol. Sci. J.*, 61, 1571–1590, 2016.

Papadopoulos V., and Giovanis D.G., Stochastic finite element methods - an introduction, *Springer*, p 138, 2018.

Papoulis, A., Probability, Random Variables, and Stochastic Processes, 3rd ed., New York: McGraw-Hill, 1991.

Pearson, K., (1930), On a New Theory of Progressive Evolution, *Annals of Eugenics*, Vol. IV, Nos. 1–2, pp. 1–40, 1930.

Pope, S.B., Turbulent Flows, *Cambridge University Press*, 2000.

Serinaldi, F., and F. Lombardo, General simulation algorithm for autocorrelated binary processes, *Phys. Rev.*, E 95, 023312, 2017a.

Serinaldi, F., and F. Lombardo, BetaBit: A fast generator of autocorrelated binary processes for geophysical research, *EPL*, 118(3), 30007, 2017b.

She, Z.S., and Leveque, E., Universal scaling laws in fully developed turbulence, *Phys. Rev. Lett.*, 72, 336, 1994.

Sklar A., Fonctions de repartition a n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Universite de Paris*, 8:229–231, 1959.

Taylor, G.I., *Proc. R. Soc. Lond.*, A 164, 476, 1938.

Tsekouras, G., and D. Koutsoyiannis, Stochastic analysis and simulation of hydrometeorological processes associated with wind and solar energy, *Renewable Energy*, 63, 624–633, 2014.

Tsoukalas I., Efstratiadis A., and Makropoulos C., Stochastic periodic autoregressive to anything (SPARTA): modelling and simulation of cyclostationary processes with arbitrary marginal distributions, *Water Resour. Res.*, 54(1):161–185, doi.org/10.1002/2017WR021394, 2018.

Tyralis, H., Koutsoyiannis, D., Simultaneous estimation of the parameters of the Hurst–Kolmogorov stochastic process, *Stoch. Environ. Res. Risk Assess.*, 25, 21–33, 2011.

Tyralis, H., D. Koutsoyiannis, and S. Kozanis, An algorithm to construct Monte Carlo confidence intervals for an arbitrary function of probability distribution parameters, *Computational Statistics*, 28 (4), 1501–1527, 2013.

Villarini, G., Mandapaka, P.V., Krajewski, W.F., Moore, R.J., Rainfall and sampling uncertainties: A rain gauge perspective, *J. Geophys. Res. Atmospheres*, 113, 2008.

Von Karman, T., The local structure of atmospheric turbulence, *Doklady Akademii Nauk. SSSR*, (67) 643, 1948.

Wilczek, M., Daitche, A. and Friedrich, R, On the velocity distribution in homogeneous isotropic turbulence: correlations and deviations from Gaussianity, *J. Fluid Mech*, 676, 191–217, 2011.

Appendix A

Here, we describe how the SMA scheme can preserve an approximation of the marginal distribution of a process through the preservation of high-order moments. Although this scheme can preserve any number of moments, here we specify the analytical solution for the preservation up to the fourth moment corresponding to kurtosis and to the fifth raw moment (for illustration). Assuming $\mathrm{E}[\underline{x}_i] = \mathrm{E}[\underline{v}] = 0$, raw moments are identical to the corresponding central moments; the p^{th} moment can be expressed through the SMA scheme as:

$$\mathrm{E}[\underline{x}_i{}^p] = \mathrm{E}\left[\left(\sum_{j=-l}^{l} a_{|j|}\underline{v}_{i+j}\right)^p\right] \tag{A-1}$$

Therefore, assuming also that $\mathrm{E}[\underline{v}^2] = 1$, the second and third raw moments can be expressed as (Koutsoyiannis, 2000):

$$\mathrm{E}[\underline{x}^2] = \left(a_0{}^2 + 2\sum_{j=1}^{l} a_j{}^2\right) \tag{A-2}$$

$$\mathrm{E}[\underline{x}^3] = \left(a_0{}^3 + 2\sum_{j=1}^{l} a_j{}^3\right)\mathrm{E}[\underline{v}^3] \tag{A-3}$$

For a raw moment of order p we use the multinomial theorem:

$$\mathrm{E}[\underline{x}_i{}^p] = \mathrm{E}\left[\left(\sum_{j=-l}^{l} a_{|j|}\underline{v}_{i+j}\right)^p\right] = \sum_{k_{-l}+k_{1-l}+\cdots+k_l=p} \binom{p}{k_{-l}, k_{1-l}, \dots, k_l} \mathrm{E}\left[\prod_{-l\le j\le l} \left(a_{|j|}\underline{v}_{i+j}\right)^{k_j}\right] \tag{A-4}$$

where $\binom{p}{k_{-l}, k_{1-l}, \dots, k_l} = \frac{p!}{k_{-l}!k_{1-l}!\dots k_l!}$, is a multinomial coefficient.

We notice that all combinations with $k_j = 1$ are zero and thus, after algebraic manipulations we obtain for $p = 4$:

$$\mathrm{E}[\underline{x}^4] = \mathrm{E}[\underline{v}^4]\sum_{j=-l}^{l} a_{|j|}{}^4 + 6\sum_{j=-l}^{l-1}\sum_{k=j+1}^{l} a_{|j|}^2 a_{|k|}^2 \tag{A-5}$$

After typical manipulations we derive the expressions for the coefficients of skewness and kurtosis shown in Eqns. (10) and (11), respectively. Also, Eqn. 11 can be further simplified for faster calculations to:

$$C_{\mathrm{k},v} = \frac{C_{\mathrm{k},x}\left({a_0}^2 + 2\sum_{j=1}^{l} {a_j}^2\right)^2 - 6\sum_{j=1}^{l} {a_j}^4 - 12a_0^2\sum_{j=1}^{l} a_j^2 - 24\sum_{j=1}^{l-1}\left(a_j^2\sum_{k=j+1}^{l} a_k^2\right)}{\left({a_0}^4 + 2\sum_{j=1}^{l} {a_j}^4\right)} \qquad \text{(A-6)}$$

with $l > 1$, while for $l = 1$ the last term of the double sum is zero.

For illustration, we also present the 5th raw moment as estimated from Eqn. A-1 and A-4 (above that moment the computation requirements highly increase due to multiplication of more than two coefficients):

$$\mathrm{E}\left[\underline{x}^5\right] = \mathrm{E}\left[\underline{v}^5\right]\sum_{j=-l}^{l} {a_{|j|}}^5 + 10\sum_{j=-l}^{l}\sum_{k=-l;k\neq j}^{l} a_{|j|}^2\, a_{|k|}^3 \qquad \text{(A-7)}$$

However, the extension to higher than the fourth moment is not required for the applications of this paper, since as illustrated in sect. 4 through the estimation of the ME distribution, the contribution of the fourth moment is small and even more so will be that of even higher moments.

Appendix B

Here, we describe how we can use selected distributions as a means to preserve the desirable statistical central moments through the SMA model. For random number generation from thin-tailed distributions we adopt an extended standardized version of the Kumaraswamy (1980) distribution (abbreviated as ESK) with probability distribution function:

$$F(x;\boldsymbol{p}) := 1 - \left(1 - \left(\frac{x-c}{d}\right)^a\right)^b \qquad \text{(B-1)}$$

where $x \in [c, c+d]$, $\boldsymbol{p} = [a, b, c, d]$, the parameters of the distribution (see also Table B-1 and B-2), with $c, d \in \mathbb{R}$ (location and scale parameters, respectively, with units same as in x) and $a, b > 0$ (dimensionless shape parameters).

Below, we estimate several statistical characteristics of the ESK distribution such as the mean, variance, and coefficients of skewness and kurtosis, as well as the minimum and maximum kurtosis as a function of skewness. A detailed analysis on the general expansion of the Kumaraswamy distribution can be found in Cordeiro and Castro (2011), and Khan et al. (2016). The ESK distribution has simple, analytical and closed expressions for its statistical central moments. Notably, we find through numerical investigation that ESK has a low kurtosis boundary based on its skewness and approximately expressed by $C_\mathrm{k} \geq {C_\mathrm{s}}^2 + 1$, which is also the mathematical boundary for the sample skewness and kurtosis (Pearson, 1930).

The central moments of the ESK distribution can be expressed as:

$$\mathrm{E}\left[\left(\underline{x}-\mu\right)^p\right] = d^p \sum_{\xi=1}^{p+1}\left((-1)^{p+1-\xi}\binom{p}{\xi-1} B_1^{p+1-\xi} B_{\xi-1}\right) \tag{B-2}$$

for $p > 1$ and where $\mu = c + dB_1$, $\binom{p}{\xi-1}$ the binomial coefficient and $B_\xi = b\mathrm{B}(1+\xi/a, b)$, with B the beta function.

Thus, the variation, skewness and kurtosis coefficients can be expressed as:

$$C_\mathrm{v} = \frac{B_2 - B_1{}^2}{(B_1 + c/d)^2},\; C_\mathrm{s} = \frac{2B_1{}^3 - 3B_1B_2 + B_3}{\left(B_2 - B_1{}^2\right)^{3/2}},\; C_\mathrm{k} = \frac{-3B_1{}^4 + 6B_1{}^2B_2 - 4B_1B_3 + B_4}{\left(B_2 - B_1{}^2\right)^2} \tag{B-3}$$

respectively. After the numerical estimation of a and b, the parameters c and d can be analytically calculated as:

$$d = \sigma / \sqrt{b\mathrm{B}\left(1+\frac{2}{a}, b\right) - b^2\mathrm{B}^2\left(1+\frac{1}{a}, b\right)},\; c = \mu - bd\mathrm{B}\left(1+\frac{1}{a}, b\right) \tag{B-4}$$

Therefore, we can use the ESK distribution to approximate a variety of thin-tailed distributions based on the estimation of a, b, c and d parameters from data. Note that if we wish to extend the SMA model to preserve additional moments, we could similarly expand the ESK distribution to simulate two (or more) additional moments, i.e., $F(x; \boldsymbol{p}) := 1 - (1 - F(x; \boldsymbol{p})^{a'})^{b'}$., with a' and b' two extra parameters.

For heavy-tailed distributions we use the standardized version of the Normal-Inverse-Gaussian (abbreviated as NIG) distribution with probability density function (cf., Barndorff-Nielsen, 1978):

$$f(x; \boldsymbol{p}) := \frac{\sqrt{a^2+b^2}\,\mathrm{e}^{b+\frac{a(x-c)}{d}}}{\pi d\sqrt{1+\left(\frac{(x-c)}{d}\right)^2}} K_1\left(\sqrt{a^2+b^2}\sqrt{1+\left((x-c)/d\right)^2}\right) \tag{B-5}$$

where $x \in \mathbb{R}$, $\boldsymbol{p} = [a, b, c, d]$, the parameters of the distribution with $c \in \mathbb{R}$, $a \neq 0$ and $b, d > 0$ (see also Table B-1 and B-2); again c, d are location and scale parameters, respectively, with units same as in x, and $a, b > 0$ are dimensionless shape parameters.

The NIG distribution has similar advantages to the ESK, such as closed expressions for the first four central moments. Also, it enables a large variety of skewness-kurtosis combinations and its random numbers can be generated almost as fast as the ESK ones through the normal variance-mean mixture:

$$x = c + \frac{a}{d} z + \sqrt{z} g \tag{B-6}$$

where

$$g{\sim}N(0,1), z{\sim}f(y;b\,,d) = d/\sqrt{2\pi y^3}\mathrm{e}^{-\frac{b^2(y/d-d/b)^2}{2y}} \tag{B-7}$$

The latter distribution is the inverse Gaussian distribution which can be easily and fast generated (e.g. Chhikara and Folks, 1989, section 4.5).

Below, we estimate the statistical characteristics of the NIG and we justify the use of the NIG distribution as a heavy-tailed distribution. Note that the central moments of the NIG function cannot be expressed as closed and analytical forms and thus, we can estimate them through the NIG characteristic function (cf. Barndorff-Nielsen, 1978):

$$\varphi_X(t) = \mathrm{E}\left[\mathrm{e}^{itX}\right] = \mathrm{e}^{ict+b-\sqrt{b^2-i2adt+d^2t^2}} \tag{B-8}$$

where the p^{th} raw moment corresponds to:

$$\mathrm{E}[X^p] = (-i)^p \lim_{t\to 0}\left(\frac{\mathrm{d}^p\varphi_X(t)}{\mathrm{d}t^p}\right). \tag{B-9}$$

Particularly, the first moment and the sequent three central moments are given by:

$$\mu = c + ad/b \tag{B-10}$$

$$\mathrm{E}\left[\left(\underline{x}-\mu\right)^2\right] = (a^2+b^2)d^2/b^3 \tag{B-11}$$

$$\mathrm{E}\left[\left(\underline{x}-\mu\right)^3\right] = \frac{3a\left((a^2+b^2)d^2/b^3\right)^{3/2}}{\sqrt{b(a^2+b^2)}} \tag{B-12}$$

$$\mathrm{E}\left[\left(\underline{x}-\mu\right)^4\right] = \frac{3\left((a^2+b^2)d^2/b^3\right)^2}{b}\left(1+\frac{4}{1+(b/a)^2}\right) + 3\left((a^2+b^2)d^2/b^3\right)^2 \tag{B-13}$$

After algebraic manipulations the coefficients of variation, skewness and kurtosis can be expressed as:

$$C_{\mathrm{v}} = \frac{a^2+b^2}{b(a+bc/d)^2}, C_{\mathrm{s}} = \frac{3a}{\sqrt{b(a^2+b^2)}}, C_{\mathrm{k}} = \frac{3}{b}\left(1+\frac{4}{1+(b/a)^2}\right)+3 \tag{B-14}$$

respectively. The NIG parameters can then be calculated from these equations as:

$$d = \frac{3\sigma\sqrt{3C_\mathrm{k}-5{C_\mathrm{s}}^2-9}}{3C_\mathrm{k}-4{C_\mathrm{s}}^2-9}, b = \frac{d}{\sigma}\sqrt{\frac{3}{C_\mathrm{k}-\frac{5}{3}{C_\mathrm{s}}^2-3}}, a = \frac{b^2 C_\mathrm{s}\sigma}{3d}, c = \mu - ad/b \tag{B-15}$$

Also, we can derive theoretically the minimum kurtosis of NIG for a given skewness:

$$C_\mathrm{k} \geq \frac{5}{3}{C_\mathrm{s}}^2 + 3 \tag{B-16}$$

with the equality holding only for the limit where the NIG tends to the normal distribution.

For the classification of tails we use the test based on the functions proposed by Klugman (1998, sect. 3.4.3; see also Halliwell, 2013) and here defined as:

$$\tau_\mathrm{r} \coloneqq -\lim_{x\to\infty}\left(\frac{\mathrm{d}f(x;\boldsymbol{p})}{f(x;\boldsymbol{p})\mathrm{d}x}\right), \tau_\mathrm{l} \coloneqq \lim_{x\to-\infty}\left(\frac{\mathrm{d}f(x;\boldsymbol{p})}{f(x;\boldsymbol{p})\mathrm{d}x}\right) \tag{B-17}$$

After calculations we get:

$$\tau_\mathrm{r} = \sqrt{a^2+b^2}/d - a/d \geq 0, \tau_\mathrm{l} = \sqrt{a^2+b^2}/d + a/d \geq 0 \tag{B-18}$$

and hence, the NIG is expected to represent a large variety of heavy-tailed distributions.

Note that again, if we wish to extend the SMA model to preserve additional moments through the NIG distribution, we could similarly expand the normal variance-mean mixture to simulate two additional moments, i.e.:

$$x = c + \frac{a}{d}z' + a'\sqrt{z'}g \tag{B-19}$$

with *a'* an extra parameter and *z'* the so-called generalized inverse Gaussian distribution.

In Fig. B-1 and B-2, we observe that the smaller possible kurtosis of the ESK distribution for a given skewness coincides with the theoretical limit defined by Pearson (1930). Also, the larger kurtosis of the ESK includes a variety of sub-Gaussian and thin-tailed distributions. On the contrary, the smaller kurtosis of the NIG distribution is very close to the larger one of the ESK and thus, it can include (as shown above) a variety of heavy-tailed distributions. Note that these two distributions are chosen to simulate unbounded processes (NIG) and bounded between two real values (ESK). We could easily find similar distributions for upper (or lower) bounded processes as, for example, the generalized Kumaraswamy or hypergeometric ones.

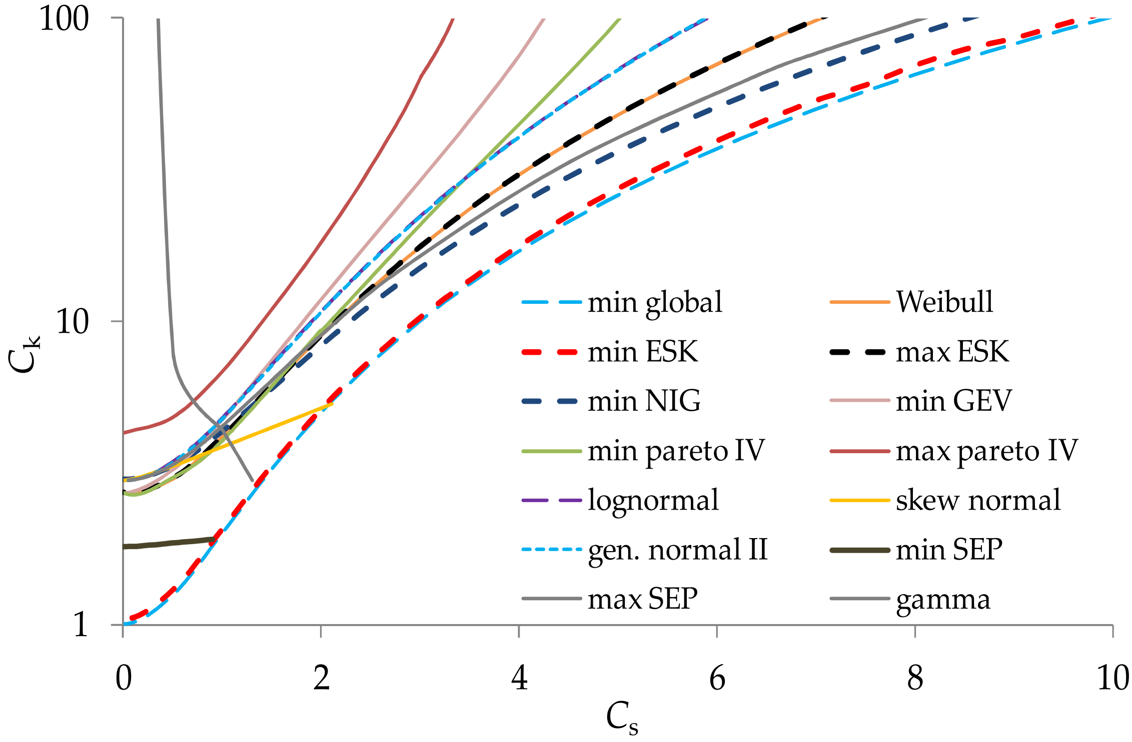

Figure B-1: Combinations of skewness and kurtosis coefficients for various two-parameter (Weibull, GEV, lognormal, generalized normal I, skew-exponential-power —SEP— and gamma), three-parameter (generalized normal II and skew normal) and the four-parameter Pareto-Burr-Fuller (PBF, further described in section 4) distribution functions along with the thin-heavy tailed separation based on the ESK and NIG functions, respectively.

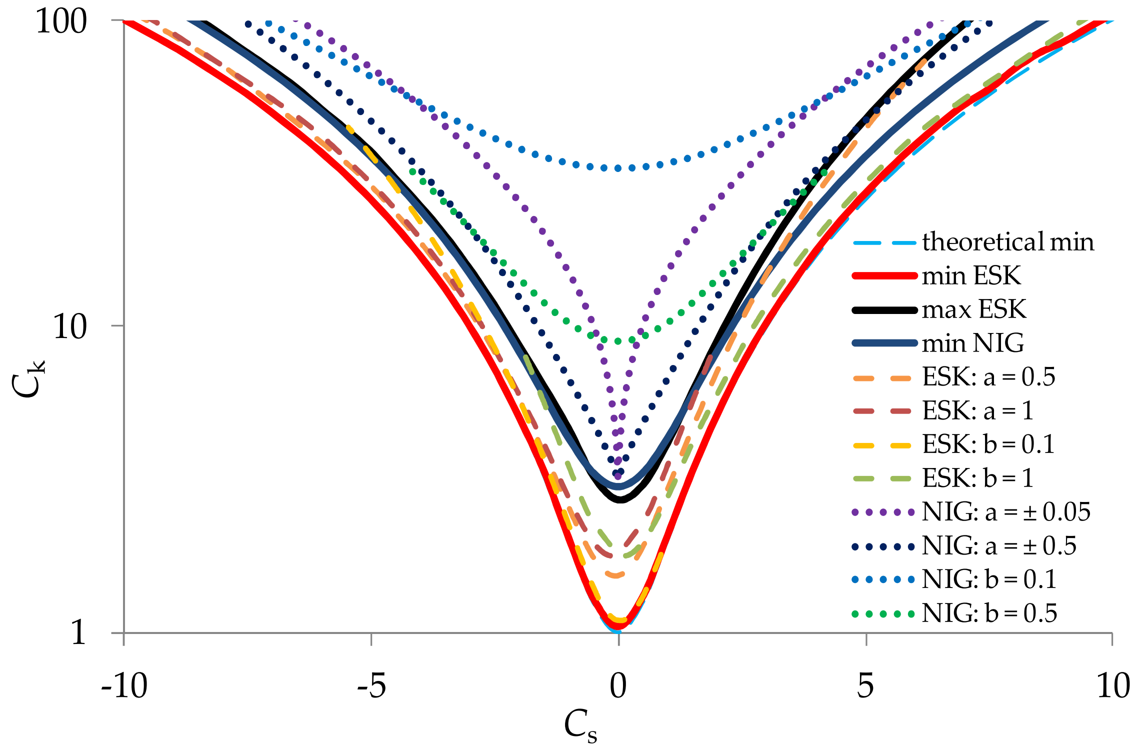

Figure B-2: Isopleths for estimated coefficients of skewness and kurtosis for the specified values of parameters *a* and *b* of the ESK and NIG distributions.

Table B-1: Mean, variance, and coefficients of skewness and kurtosis for the ESK and NIG distributions. Note that $B_i = b\mathrm{B}(1 + i/a, b)$, where $\mathrm{B}(x, y)$ is the beta function and i an integer.

	ESK	NIG
μ	$c + dB_1$	$c + ad/b$
σ^2	$d^2(B_2 - {B_1}^2)$	$\frac{(a^2 + b^2)d^2}{b^3}$
C_s	$\frac{2{B_1}^3 - 3B_1B_2 + B_3}{(B_2 - {B_1}^2)^{3/2}}$	$\frac{3a}{\sqrt{b(a^2 + b^2)}}$
C_k	$\frac{-3{B_1}^4 + 6{B_1}^2B_2 - 4B_1B_3 + B_4}{(B_2 - {B_1}^2)^2}$	$\frac{3}{b}\left(1 + \frac{4}{1 + (b/a)^2}\right) + 3$
$\min C_\mathrm{k}$	$\approx {C_\mathrm{s}}^2 + 1$	$= \frac{5}{3}{C_\mathrm{s}}^2 + 3$
$\max C_\mathrm{k}$	$\approx \frac{5}{3}{C_\mathrm{s}}^2 + 3$ *	$+\infty$

* This is a fair approximation only for $C_\mathrm{s} \le -2$. A more exact but empirical approximation for $-10 \le C_\mathrm{s} \le 10$, can be given by: $0.039{C_\mathrm{s}}^3 + 1.724{C_\mathrm{s}}^3 + 0.032{C_\mathrm{s}}^3 + 2.7$. Note that the maximum kurtosis for the ESK for a given skewness approximately coincides with the kurtosis of the Weibull distribution (Fig. B-1).

Table B-2: Parameters of the ESK and NIG distributions in terms of the mean, standard deviation, and coefficients of skewness and kurtosis (see also Fig. B-2).

distribution	ESK	NIG
a	non-analytical *	$\frac{b^2 C_\mathrm{s}\sigma}{3d}$
b	non-analytical *	$\frac{d\sqrt{3}}{\sigma\sqrt{C_\mathrm{k} - \frac{5}{3}{C_\mathrm{s}}^2 - 3}}$
c	$\mu - dB_1$	$\mu - ad/b$
d	$\frac{\sigma}{\sqrt{(B_2 - {B_1}^2)}}$	$\frac{3\sigma\sqrt{3C_\mathrm{k} - 5{C_\mathrm{s}}^2 - 9}}{3C_\mathrm{k} - 4{C_\mathrm{s}}^2 - 9}$

*The two parameters of the ESK distribution a and b can be estimated by solving numerically the two following equations: $C_\mathrm{s} = (2{B_1}^3 - 3B_1B_2 + B_3)/(B_2 - {B_1}^2)^{3/2}$ and $C_\mathrm{k} = (-3{B_1}^4 + 6{B_1}^2B_2 - 4B_1B_3 + B_4)/(B_2 - {B_1}^2)^2$.

In case additional moments need to be preserved, a more generalized methodology includes the use of the ME distribution, which can be applied for any type of distribution (see Eqn. 12-13):

$$f(x;\boldsymbol{\lambda}):=\frac{1}{\lambda_0}\mathrm{e}^{-\left(\frac{x}{\lambda_1}+\operatorname{sign}(\lambda_2)\left(\frac{x}{\lambda_2}\right)^2+\left(\frac{x}{\lambda_3}\right)^3+\operatorname{sign}(\lambda_4)\left(\frac{x}{\lambda_4}\right)^4+\left(\frac{x}{\lambda_3}\right)^5+\operatorname{sign}(\lambda_6)\left(\frac{x}{\lambda_6}\right)^6+\cdots+\left(\frac{x}{\lambda_m}\right)^m\right)} \quad \text{(B-20)}$$

where $\boldsymbol{\lambda}= [\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \dots, \lambda_m]$, (where m even) with $\lambda_0,\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \dots, \lambda_m$ having same units as x and with $m + 1$ constraints, where m is the number of moments we wish to preserve:

$$\int_{-\infty}^{\infty} x^r f(x;\boldsymbol{\lambda})\mathrm{d}x = \mathrm{E}\left[\underline{x}^r\right], \text{for } r = 0, \dots, m \quad \text{(B-21)}$$

For the generation scheme of the above distribution, we may use the random number generator described in the next steps. After we estimate the $\boldsymbol{\lambda}$ parameters of the MED we can rewrite the above equation as:

$$f'(x;\boldsymbol{\lambda}'):=\lambda'_0\mathrm{e}^{-\left((\lambda'_1x+\lambda'_2)^2+(\lambda'_3x+\lambda'_4)^4+(\lambda'_5x+\lambda'_6)^6+\cdots\right)} \quad \text{(B-22)}$$

with exactly the same number of unknown parameters (note that an exact solution for $\boldsymbol{\lambda}'$ always exists).

After we estimate the new parameters $\boldsymbol{\lambda}'$ we can approximate the above distributions with an auxiliary distribution function for an even m:

$$g(x;\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}) = \begin{cases} c_1\mathrm{e}^{-(a_1x+b_1)^2}, d_1 < x < d_2 \\ c_2\mathrm{e}^{-(a_2x+b_2)^4}, d_3 < x \le d_1, d_2 \le x < d_4 \\ c_3\mathrm{e}^{-(a_3x+b_3)^6}, d_5 < x \le d_3, d_4 \le x < d_6 \\ \dots \end{cases} \quad \text{(B-23)}$$

where $\boldsymbol{a}= [a_1, a_2, a_3, \dots]$, $\boldsymbol{b}= [b_1, b_2, b_3, \dots]$, $\boldsymbol{c}= [c_1, c_2, c_3, \dots]$ and $\boldsymbol{d}= [d_1, d_2, d_3, \dots]$ with the last two d parameters to be -∞ and +∞, respectively. The above distribution is subjected to the constraint $\int_{-\infty}^{\infty} g(x;\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d})\mathrm{d}x = 1$, and continuity in all its branches.

After the estimation of all the parameters through optimization techniques (so as g to be as close as possible to f') we can use the rejection method (Papoulis, 1990, pp. 261-263) to generate random number for the MED. Note that for the generation of each branch function of the g distribution, we can use the random number generator of the powered-exponential distribution function through the generation of the gamma distribution function (the latter can be also generated using the rejection method as described in Koutsoyiannis and Manetas, 1996).

Appendix C

Here we describe how the SMA model can be used to cope with non-stationary processes. The general idea is to convert non-stationary processes to stationary ones, so that eventually the simulation is made for a stationary process (Dimitriadis, 2017). This conversion is achieved by appropriate transformations or by separating them into segments, as for example in the case of cyclostationary processes. While in the recent literature there is no shortage of publications seeking or assuming non-stationarity, this may just reflect incomplete understanding of what stationarity is (Koutsoyiannis and Montanari, 2015). A common confusion is that non-stationarity is regarded as a property of the natural process, while in fact it is a property of a mathematical (stochastic) process. In non-stationary processes some of the statistical properties change in time in a deterministic manner. The deterministic function describing the change in the statistical properties is rarely known in advance and, in studies claiming non-stationarity, is typically inferred from the data. However, it is impractical or even impossible to properly fit a non-stationary mathematical process to time series, as in nature only one time series of observations of a certain process is possible, while the definition of stationarity or non-stationarity relies on the notion of an ensemble of time series.

A simple example of how we can deal with a non-stationary process through a stationary one follows. We consider an HK process (denoted as $\underline{x}$) with H = 0.8 μ = 0 and σ = 1 and by aggregation we also take the cumulative process (denoted as $\underline{y}$, i.e. $\underline{y}_i = \underline{y}_{i-1} + \underline{x}_i$). Figure C-1 shows a time series generated from $\underline{x}$ and the corresponding time series from $\underline{y}$. Clearly, $\underline{x}$ is stationary and $\underline{y}$ is non-stationary (the so-called fractional Brownian noise). If we have the information about the theoretical basis of the two processes, then it is trivial to correctly model them (Koutsoyiannis, 2016). In particular, we will know that the mean of the process $\underline{y}$ is constant (zero, not a function of time) while its variance is an increasing function of time (a power-law function of i). Otherwise, if the only available information is the time series of $\underline{y}$, then we may be tempted to assume a linear trend for the mean of $\underline{y}$ and express the mean of the process as a linear function of time, $\underline{\mu}_i = a\,\underline{i} + b$ (with a and b the parameters of the slope and intercept of a regression line on the time series). This, however, would be plain wrong as in fact (by construction) the mean of $\underline{y}$ is zero for any time i. In addition, the introduction of the two extra parameters (i.e., $\underline{a}$ and b) has negative implications in terms of the overall uncertainty of the model, which would cease to be parsimonious. But again, even with this wrong assumption, the next step would be to construct a stationary model, i.e., $\underline{z}_i = \underline{y}_i - a\,\underline{i} - b$ and use that model in simulations. The correct approach for this case would be to construct the time series of $\underline{x}$ by differentiation of $\underline{y}$ (i.e., $\underline{x}_i = \underline{y}_i - \underline{y}_{i-1}$), which is stationary, and use the stationary process $\underline{x}$ for stochastic simulation; then a synthetic time series of the non-stationary process $\underline{y}$ will be constructed from a time series of $\underline{x}$. Thus, in all cases, whether with correct or incorrect assumptions, the stochastic simulation is always done for a stationary process.

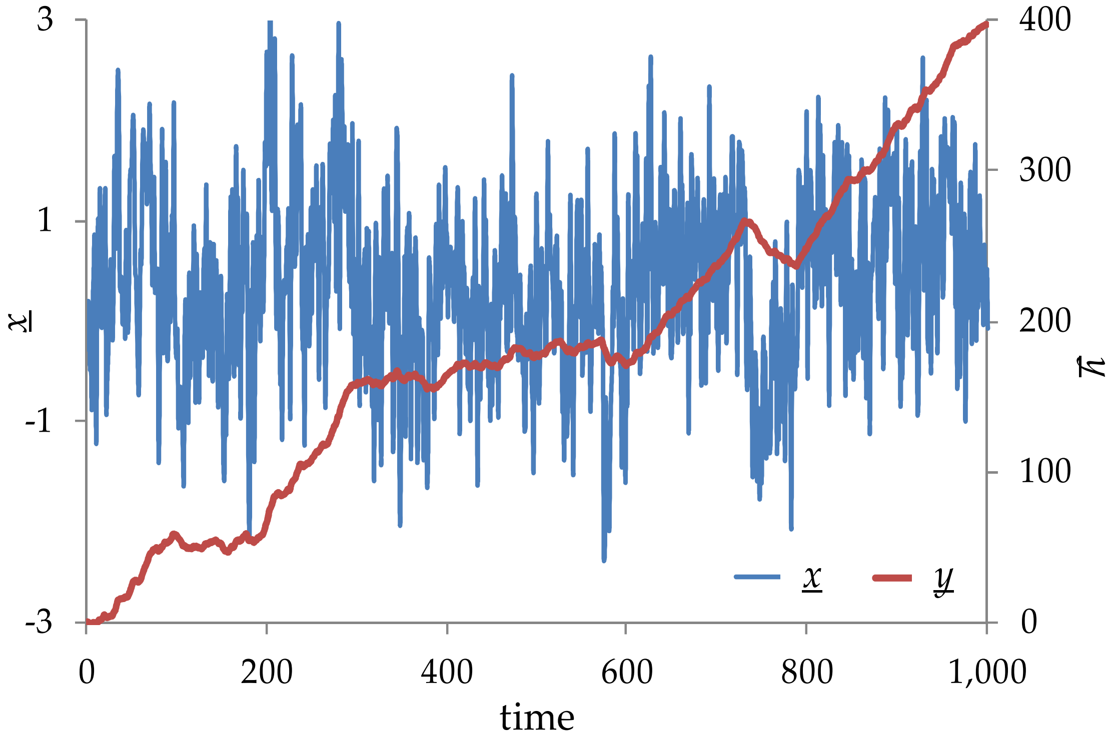

Figure C-1: Time series with length 1000 from the example processes $\underline{x}$ and $\underline{y}$.

Appendix D

The first applied generic schemes for a stochastic synthesis are the implicit ones, i.e. those approximating the distribution and dependence structure of a process through non-linear transformations. These non-linear transformations (or else known as copula, Hoeffding, 1940; Frechet, 1951; Sklar, 1959; Nelsen, 2006, and references therein) are often based on the autocovariance function for any distribution function, where the uniform is usually preferred for reasons of simplicity, whereas for reasons of flexibility the Gaussian distribution (the so-called Gaussian-copula; Lebrun and Dutfoy, 2009) can be also used (for the bivariate Gaussian copula see Nataf, 1962; Serinaldi and Lombardo, 2017a, Tsoukalas et al., 2018; Papadopoulos and Giovanis, 2018; and references therein). This scheme is also known as Nataf transformation, but here we propose to use the name Hoeffding-Frechet-Sklar-Nataf (HFSN) transformation, since Nataf wrote a half-page conference paper (presented by Frechet) just mentioning (without further analyzing it) a specific case of the general methodology earlier discussed by Hoeffding, Frechet and Sklar. The general HFSN transformation can be written as:

$$\begin{aligned}\rho_{x_i x_j}\sigma^2 + \mu^2 &= \mathrm{E}[\underline{x}_i \underline{x}_j] = \mathrm{E}\left[T\left(\underline{y}_i\right) T\left(\underline{y}_j\right)\right] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T(y_i)T(y_j) f\left(y_i, y_j; \rho_{y_i y_j}\right) \mathrm{d}y_i\, \mathrm{d}y_j\end{aligned} \qquad \text{(D-1)}$$

where $\rho_{x_i x_j}$ and $\rho_{y_i y_j}$ are the cross-correlations between x_i and x_j as well as y_i and y_j, respectively; μ and σ are the process mean and standard deviation; $f(y_i, y_j; \rho_{y_i y_j})$ is the joint distribution between y_i and y_j; $T(y_i)$ and $T(y_j)$ are the transformations of the original known distribution of x_i and x_j to the selected distribution of y_i and y_j, respectively (see below for an example of such transformations). In case that $\underline{y}$ is for example N(0,1) distributed, the bivariate N(0,1) is used, i.e. $f(y_i, y_j; \rho_{y_i y_j}) = \mathrm{e}^{-1/2(y_i^2+y_j^2-2y_i y_j \rho_{y_i y_j})/(1-\rho_{y_i y_j}^2)}/(2\pi(1-\rho_{y_i y_j}^2)^{1/2})$.

Similar implicit schemes are developed based on the power spectrum (Cugar, 1968; Lavergnat, 2016 and references therein). This implicit scheme can be also introduced through the climacogram, i.e.:

$$
\begin{aligned}
\gamma_x(k) + \mu^2 &= \mathrm{E}\left[\left(\underline{x}^{(k)}\right)^2\right] = \mathrm{E}\left[\left(T\left(\underline{y}\right)^{(k)}\right)^2\right] \\
&= \frac{1}{k^2}\int_{-\infty}^{+\infty}\left(\int_0^k T\big(y(t)\big)\mathrm{d}t\right)^2 f^{(k)}\left(y; \gamma_y(k)\right)\mathrm{d}y
\end{aligned}
\tag{D-2}
$$

where $\underline{x}^{(k)} \coloneqq \frac{1}{k}\int_0^k \underline{x}(t)\mathrm{d}t$ is the averaged process of the continuous-time process $\underline{x}(t)$, μ is the process mean, and $T(\underline{y})$ is a transformation function of the original process $\underline{y}$ (with the selected uniform, Gaussian etc. distribution function and with an unknown γ_y dependence structure) to the desired one $\underline{x}$ (with known density distribution $f(x)$ and dependence structure γ_x adjusted for bias). For example, a $\underline{y} \sim$ N(0,1) can be easily transformed to a $\underline{x} \sim F(x) = 1 - (a/x)^b$ by $\underline{x}(t) = T(\underline{y}(t)) = a/(1-(1/2(1+\mathrm{erf}\,(\underline{y}(t)/\sqrt{2})))^{1/b})$.

The climacogram-implicit scheme has been applied to several (stationary and single/double cyclostationary) processes, such as solar radiation (Koudouris et al., 2017), wave height and wind process for renewable energy production (Moschos et al., 2017), as well as for the wind speed using a special case of the PBF distribution (Deligiannis et al., 2016) but also a generalized non-linear transformation (equivalent to a distribution function) based on the maximization of entropy when the distribution function is unknown (Dimitriadis and Koutsoyiannis, 2015b). Note that in all the above applications the same dependence structure is used for the original and the transformed process, since a small deviation between them is noticed and therefore, additional trials are considered not necessary.

A difficulty with the implicit schemes is that they involve non-linear transformations and double integration (both of which may highly increase the numerical burden, even though fast algorithms have been discussed by Serinaldi and Lombardo, 2017b). Several exact solutions of the implicit scheme may exist even though an exact solution may not be possible for some processes (especially in very strong correlation structure as is the case in the small scales related to the fractal behaviour). Furthermore, there is no guarantee that the resulting autocorrelation structure of the transformed process will be symmetric positive definite (Lebrun and Dutfoy, 2009). In addition to the above, the transformation cannot be invariant with respect to the time lag or time scale, while the fractal and HK behaviour cannot be easily handled since the transformation is invariant with respect to the zero and infinite time scale. Some of these limitations can be dealt with through a cautiously constructed binary scheme, a multivariate Gaussian (or with other distribution such as the uniform one) copula scheme, a Monte-Carlo approach to identify the unknown dependence structure, or a properly handled disaggregation scheme for generating events of the process or more generally, by adjusting any desired stochastic properties (dependence structure and distribution function) to each scale (Lombardo et al., 2017).

However, three of these problems concerning the implicit schemes can be easily dealt by the proposed explicit scheme. Namely, these are (a) the inability of simulating the effect of the fractal behaviour of a process at small scales, in which the correlation structure is very strong, (b) the difficulties in preserving long-term persistence, and in particular its variability (see below) and (c) the effect of the statistical bias (Dimitriadis, 2017, sect. 2.4.5). In the first two problems the implicit schemes simulate the fractal and HK behaviour and bias of the non-linear transformation process $T(\underline{y})$, i.e. $f\left(T(y_i), T(y_j); \rho_{T(y_i)T(y_j)}\right)$, which due to discretization and finite length are not equal to the ones of the infinite length size continuous-time process, i.e. $f\left(y_i, y_j; \rho_{y_i y_j}\right)$. Because of the theoretical origins of these two limitations the implicit schemes are only theoretically valid for processes with short-term persistence and no fractal behaviour (i.e., H = M = 0.5), and can be used for long-term processes with fractal behaviour only as a rough approximation.

For illustration, we present a simple example to highlight the above problems related to the implicit schemes such as the simplest case of the HFSN transformation. Particularly, we generate (through the SMA scheme) data from a N(0,1) distribution with an HHK dependence structure (q = 10, M = 1/3, H = 5/6) and we transform them to a Pareto II distribution (a = b = 10) through its inverse distribution function. Subsequently, we estimate (through Monte-Carlo techniques) the expected climacogram of the transformed process and we perform separately a sensitivity analysis (as in Dimitriadis et al., 2016a) for the original (Gaussian-HHK) and transformed Pareto-HHK (q = 6.538, M = 0.431, H = 0.832) processes (Figure D-1). Furthermore, we simulate the same transformed process but now using the explicit scheme proposed in this paper with just four moments (Figure D-1). Finally, we compare the differences between the two methods in the simulation of the fractal and HK behaviour as well as of the distribution function. We observe that the variances of the sample variance of the two schemes are very different (although their mean values coincide) and that the implicit scheme overestimates it (by a factor of 10). Note that the true variance of the sample variance corresponds to that of the explicit scheme, since it explicitly preserves both the climacogram and the coefficient of kurtosis and thus, can approximate all the arising moments through the SMA scheme, such as regular moments (i.e. $E[X]$,$E[X^2]$ and $E[X^4]$) but also joint moments (i.e. $E[X_i X_j^2]$ and $E[X_i^2 X_j^2]$), which are all function of a combination of the SMA weight coefficients and the marginal moments of the white noise process (that are both exactly preserved through the explicit scheme). We believe that the reason why this is not identified in some of the recent literature is probably because all applications of the second-order implicit schemes are based solely on the preservation of the expected (mean) value of the dependence structure and not additionally on its variance (or its distribution in general) for each scale. This can be dealt by higher-order (more than two) copula schemes (or by selecting other distributions, such as the uniform one, for the transformation) but some difficulties may still remain (see above for three major ones).

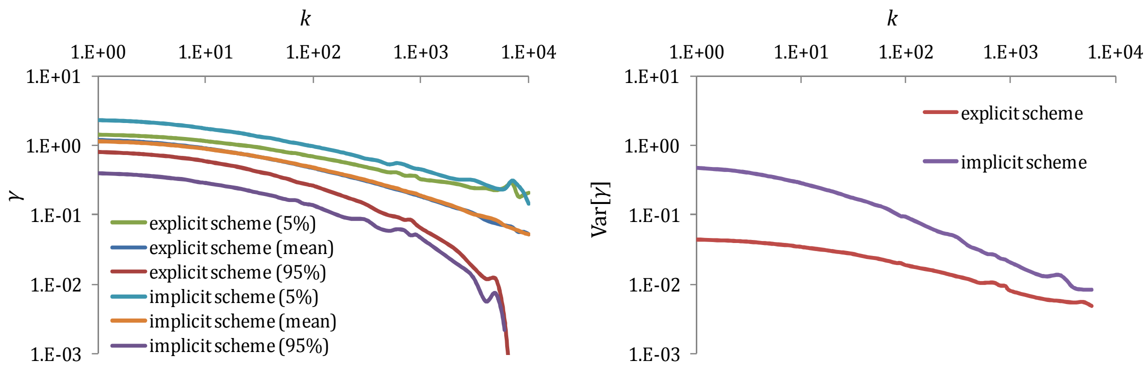

Figure D-1: Mean, 5%, 95% [left] and variance [righr] of the sample climacogram for the implicit and explicit (preserving four moments) scheme of a Pareto II (a = b = 10) and HHK (q = 6.538, M = 0.431, H = 0.832) process.