Article Towards generic simulation for demanding stochastic processes

Demetris Koutsoyiannis and Panayiotis Dimitriadis

Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, 15780 Athens, Greece Correspondence: dk@itia.ntua.gr

Abstract: We outline and test a new methodology for genuine simulation of stochastic processes with any dependence and any marginal distribution. We reproduce time dependence with a generalized, time symmetric or asymmetric, moving-average scheme. This implements linear filtering of non-Gaussian white noise, with the weights of the filter determined by analytical equations in terms of the autocovariance of the process. We approximate the marginal distribution of the process, irrespective of its type, using a number of its cumulants, which in turn determine the cumulants of white noise in a manner that can readily support the generation of random numbers from that approximation, so that it be applicable for stochastic simulation. The simulation method is genuine as it uses the process of interest directly without any transformation (e.g. normalization). We illustrate the method in a number of synthetic and real-world applications with either persistence or antipersistence, and with non-Gaussian marginal distributions that are bounded, thus making the problem more demanding. These include distributions bounded from both sides, such as uniform, and bounded form below, such as exponential and Pareto, possibly having a discontinuity at the origin (intermittence). All examples studied show the satisfactory performance of the method.

Keywords: stochastics; stochastic processes; stochastic simulation; Monte Carlo simulation; long range dependence; persistence; Hurst-Kolmogorov dynamics; climacogram; cumulants; intermittence

1. Introduction

Reviews of the historical evolution of simulation of stochastic processes with its different schools has recently been provided by Koutsoyiannis [1] and Beven [2]. In most scientific disciplines the dominant methods are those of the so-called Time Series School with families of models known with the acronym ARMA (standing for autoregressive moving average) and also called Box-Jenkins models (owing to the influential book of these authors [3], thus confirming Stigler's law of eponymy [4], because in fact they were introduced earlier by Whittle [5-7]). Despite their popularity, these models have several problems, such as their lack of parsimony (except for the simplest of them), the inability to model long-range dependence (LRD), as well as to simulate non-Gaussian processes. On the other hand, both these features are profoundly present in most geophysical processes [8]. An extension of these models applicable to processes with LRD was proposed by Hosking [9] under the acronym ARFIMA (with the letter 'F' standing for fractional differencing and the letter 'I' for integrated). Again, these are good for Gaussian processes. Koutsoyiannis (2000) [10] introduced the symmetric moving average (SMA) scheme to replace ARMA models with a generic approach (more recently advanced in [11]) capable of reproducing any aspect of time dependence, short-range (SRD) or long-range (LRD). This scheme can also preserve the skewness of non-Gaussian processes but has difficulties in dealing with higher-order moments and particularly with strongly intermittent processes, such as rainfall at small time scales.

For the latter, point process (clustered) models have been devised [12-15]. An advantage of these types of models is the mechanistic representation of certain aspects of the process, such as the arrival and cease of a storm event. The disadvantages are mainly focused on the preservation of the dependence structure at multiple scales and their difficulty in application in multivariate or multiscale schemes. For this reason, Koutsoyiannis et al. [16], even though they used a 3D extension of a point process model (the so-called Gaussian displacement spatial-temporal rainfall [17]), resorted to a linear generation scheme for an application to multivariate rainfall disaggregation.

Several other modelling schemes use transformations of the process of interest, mostly within a copula context [18-21], with the most widely applied transformation resulting in a Gaussian process (normalization) [22,23]. However, such transformation schemes inherit some of the limitations of the parent process. For example, it is well-known that a Gaussian process is necessarily symmetric in time and thus cannot capture time directionality, the so-called time's arrow [24-26], while it is known that in several natural processes time's arrow is present [24,27]. On the other hand, processes with asymmetric distributions can also exhibit asymmetry in time.

A more general algorithm for generation of any type of marginal distribution has recently been proposed by Lombardo et al. [28] but only under the condition of Markov dependence, thus leaving out problems with more complex dependence, including LRD.

For these reasons, it is necessary to develop *genuine stochastic simulation* procedures which will be able to generate non-Gaussian processes without transformations to a Gaussian or other distribution. Such procedures have already been discussed in earlier works, referring to the explicit preservation of four moments in a time-symmetric setting [29] as well as preservation of distributions in terms of cumulants, rather than moments [30,31]. However, the general idea of the latter works has never been applied in practice to test its effectiveness. This is the subject of this paper. In the following sections we outline the new methodology for genuine simulation procedures with any dependence and any marginal distribution (section 2), and illustrate it in a number of synthetic and real-world applications (section 3). In addition, we study the problem of approximating any distribution, if a number of its cumulants are known, in a manner that can readily support the generation of random numbers from that approximation, so that the approximation be applicable for stochastic simulation.

The simulation model developed is a linear stochastic model. As nonlinearity is fashionable, some may think that the linearity of the approach proposed is a limitation or even a severe drawback. The reality however is different because linearity and nonlinearity have different meaning in deterministic and stochastic approaches. In the latter, linearity is a powerful characteristic, enabling its extension in demanding problems, such as multivariate models and coupling of models of different temporal or spatial scales [32] (also known as downscaling or disaggregation). In this respect it is relevant to recall the notion of Wold decomposition of stochastic processes. Specifically, Wold [33,34] proved that any stochastic process (even though he referred to it as a time series) can be decomposed into a regular process (i.e., a process linearly equivalent to a white noise process) and a predictable process (i.e., a process that can be expressed in terms of its past values). Thus, nonlinearity is relevant to the predictable part, as this is purely deterministic, while for the regular part linearity suffices.

2. Methods

2.1 Preliminaries

We denote \underline{x} a stochastic (random) variable, underlining its symbol in order to distinguish it from a regular variable, $F(x) \coloneqq P\{\underline{x} \le x\}$ its probability distribution function, $\overline{F}(x) \coloneqq 1 - F(x) = P\{x > x\}$ its tail function (probability of exceedance) and $f(x) \coloneqq$ df(x)/dx its density function. Furthermore, we denote $\underline{x}(t)$ a stochastic process at continuous time t (i.e., a family of stochastic variables \underline{x} indexed by time t) and $\underline{x}_{\tau} \coloneqq \frac{1}{D} \int_{(\tau-1)D}^{\tau D} \underline{x}(t) dt$ its discrete time representation at equidistant times with temporal resolution D, i.e., $t_{\tau} = \tau D$, for an integer τ . In a discrete-time stochastic process it is convenient to define the *return period*, T, of the event $\{\underline{x}_{\tau} \leq x\}$ as the average time between two occurrences of the event. It is shown [31] that the following relationship holds true for any stochastic process (irrespective of time dependence):

$$\frac{T(x)}{D} = \frac{1}{\overline{F}(x)} \tag{1}$$

In other words, this one-to-one correspondence allows the return period to be used in place of the tail function or the distribution function in several applications (e.g. in probability plots); this has been the case for many years, particularly in engineering applications.

2.2 Moments and cumulants

The expectation of any function $g(\underline{x})$ of the stochastic variable \underline{x} is defined as:

$$\mathbf{E}[g(\underline{x})] \coloneqq \int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x \tag{2}$$

where we remind that $g(\underline{x})$ is a stochastic variable per se. For $g(\underline{x}) = \underline{x}^p$, we get the *non-central moment* of order *p* (or *pth raw moment* or *pth moment about the origin*):

$$\mu'_p \coloneqq \mathbf{E}[\underline{x}^p] \tag{3}$$

with the particular case p = 1 defining the mean:

$$\iota \coloneqq \mu_1' = \mathbb{E}[\underline{x}] \tag{4}$$

The central moment of order *p* is the expectation of $g(\underline{x}) = (\underline{x} - \mu)^p$:

$$\mu_p \coloneqq \mathrm{E}[(\underline{x} - \mu)^p] \tag{5}$$

with the particular case p = 2 defining the variance:

$$\mu_2 \equiv \gamma \coloneqq \mathrm{E}[(\underline{x} - \mu)^2] \coloneqq \sigma^2 \tag{6}$$

where its square root σ is the standard deviation.

By choosing $g(\underline{x}) = e^{t\underline{x}}$ for any *t*, the logarithm of the resulting expectation is called the *cumulant generating function*:

$$X(t) \coloneqq \ln \mathbb{E}[e^{t\underline{x}}] \tag{7}$$

The power series expansion of the cumulant generating function i.e.:

$$K(t) = \sum_{p=1}^{\infty} \kappa_p \frac{t^p}{p!}$$
(8)

defines the *cumulants* κ_p . It is noted that cumulants were introduced by Thielle as early as in 1889 [35] and refined in 1899 ([36], translated to English by Hald [37]) under the name *half-invariants*. The name *cumulants* was first used by Fisher [38] at the suggestion of Hotelling [39].

Cumulants are related to noncentral moments of same and lower order by:

$$\mu'_{p} = \sum_{i=0}^{p-1} {p-1 \choose i} \kappa_{p-i} \mu'_{i}, \qquad \kappa_{p} = \mu'_{p} - \sum_{i=1}^{p-1} {p-1 \choose i} \kappa_{p-i} \mu'_{i}$$
(9)

with μ'_0 = 1. A simple proof of these equations has been provided by Smith (1995) [40], but the recursive relationships had been already implied by Thielle [35,37]. Note that Equation

(9) links cumulants with noncentral moments. The relationship of cumulants with central moments is generally more complex, but for small *p* it takes the following simple forms:

$$\kappa_0 = \mu_1 = 0, \quad \kappa_1 = \mu'_1 = \mu, \quad \kappa_2 = \mu_2, \quad \kappa_3 = \mu_3, \quad \kappa_4 = \mu_4 - 3\mu_2^2$$
(10)

Equation (9) is very powerful as it allows simple calculation of cumulants from noncentral moments and vice versa in a recursive manner. Specifically, we notice that for the calculation of the moment or the cumulant of order p, the sums appearing in Equation (9) contain terms of order smaller than p.

The importance of cumulants results from their homogeneity and additivity properties, which are shown in Table 1. Most importantly, for a stochastic variable that is the linear combination (weighted sum) of r independent variables, the cumulants of the resultant are also a linear combination of the cumulants of the constituents. On the other hand, application of conditioning, also contained in Table 1, is similarly useful as it allows simulation of distributions that are mixtures of other distributions or have discontinuities in their distribution functions. As shown in Table 1, the effect of conditioning is more easily expressed in terms of moments, but Equation (9) readily allows the subsequent evaluation of cumulants.

Table 1. Typical operations useful in simulation and their mathematical handling.

Operation	Mathematical relationship	Eqn. no.
Shift of origin	$\kappa_p[\underline{x} + c] = \begin{cases} \kappa_1[\underline{x}] + c & p = 1\\ \kappa_p[\underline{x}] & p > 1 \end{cases}$	(11)
Multiplication by a constant (<i>a</i>)	$\kappa_p[\underline{a\underline{x}}] = \kappa_p[\underline{x}] a^p$	(12)
Linear combination of independent variables	$\kappa_p[a_1\underline{x}_1 + \dots + a_r\underline{x}_r] = a_1^p \kappa_p[\underline{x}_1] + \dots + a_r^p \kappa_p[\underline{x}_r]$	(13)
Conditioning on an event A_1 with probability $P_1 \coloneqq P(A_1)$, where the complementary event A_2 has probability $1 - P_1 = P(A_2)$	$\mu_p'[\underline{x}] = P_1 \mu_p'[\underline{x} A_1] + (1 - P_1)\mu_p'[\underline{x} A_2]$	(14)
Conditioning on an event A_1 with probability $P_1 \coloneqq P(A_1)$, where $\underline{x} = c$ (constant) upon the complementary event A_2	$\mu_p'[\underline{x}] = P_1 \mu_p'[\underline{x} A_1] + (1 - P_1)c^p$	(15)
Conditioning on an event A_1 with probability $P_1 := P(A_1)$, where $\underline{x} = 0$ upon the complementary event A_2	$\mu_p'[\underline{x}] = P_1 \mu_p'[\underline{x} A]$	(16)

All common distribution functions used in a wide range of stochastic applications have elegant analytical expressions either for their moments or the cumulants of any order, and in some cases of both. These are gathered in Table 2 for distributions with finite domain, in Table 3 for distributions with infinite domain but with all their moments finite, and in Table 4 for the heavy-tailed distributions with upper-tail index ξ ; in the latter case, both moments and cumulants exist for $p < 1/\xi$ and are infinite for larger p. The following notes apply to these tables:

1. The meaning of the parameters is the following.

(a) Dimensional parameters, with dimensions identical to those of the stochastic variable \underline{x} : μ : mean; $\sigma > 0$: standard deviation; $\lambda > 0$: scale parameter; a, b: lower and upper bound of \underline{x} .

(b) Dimensionless parameters: $\xi > 0$: upper-tail index; $\zeta > 0$: lower-tail index; $\varsigma > 0$: additional shape parameter, $P_i \in [0,1]$: probability.

5 of 24

Name, domain	Probability density or distribution function	Moments, μ'_p	Cumulants, κ _p
Impulse, $\underline{x} = \mu$	$f(x) = \delta(x-\mu)$	μ^p	$\begin{cases} \mu & p = 1 \\ 0 & p > 1 \end{cases}$
Finite number of impulses, $\underline{x} \in \{x_1, \dots, x_n\}$	$f(x) = \sum_{i=1}^{n} P_i \delta(x - x_i)$	$\sum_{i=1}^{n} P_i x_i^p$	
Uniform, $a \leq \underline{x} \leq b$	$f(x) = \frac{1}{b-a}$	$\frac{b^{p+1} - a^{p+1}}{(p+1)(b-a)}$	$\begin{cases} \mu_1' = \frac{a+b}{2} & p = 1\\ \frac{(b-a)^p B_p}{p} & p \text{ odd}\\ 0 & p \text{ even} \end{cases}$
Beta, $0 \le \underline{x} \le b$	$f(x) = \frac{\left(\frac{x}{b}\right)^{\zeta-1} \left(1 - \frac{x}{b}\right)^{\zeta-1}}{B\left(\zeta, \varsigma\right)}$	$\frac{\Gamma\left(\zeta+\varsigma\right)\Gamma\left(p+\zeta\right)}{\Gamma\left(\zeta\right)\Gamma\left(p+\zeta+\varsigma\right)}b^{p}$	
Kumaraswamy, $0 \le \underline{x} \le b$	$F(x) = 1 - \left(1 - \left(\frac{x}{b}\right)^{\zeta}\right)^{\zeta}$	$\zeta B\left(\zeta, \frac{p+\zeta}{\zeta}\right)b^p$	

Table 2. Noncentral moments and cumulants of common distributions with finite domain (all moments and cumulants exist).

Table 3. Noncentral moments and cumulants of common distributions with zero upper-tail index (all moments and cumulants exist).

Name, domain	Probability density or distribution function	Moments, μ'_p	Cumulants, κ_p
Poisson $\underline{x} = j, j \in \mathbb{N}_0$	$f(x) = e^{-\varsigma} \sum_{i=1}^{\infty} \frac{\varsigma^{j}}{j!} \delta(x-k)$		ς
Exponential, $x \ge 0$	$f(x) = \mathrm{e}^{-x/\mu}/\mu$	$p! \mu^p$	$(p-1)! \mu^p$
Gamma, $\underline{x} \ge 0$	$f(x) = \frac{(x/\lambda)^{\zeta - 1} e^{-x/\lambda}}{\lambda \Gamma(\zeta)}$	$\frac{\Gamma(p+\zeta)}{\Gamma(\zeta)}\lambda^p$	$\zeta(p-1)! \lambda^p$
Generalized gamma, $\underline{x} \ge 0$	$f(x) = \frac{1}{\lambda \Gamma(\zeta/\zeta')} \left(\frac{x}{\lambda}\right)^{\zeta-1} \exp\left(-\left(\frac{x}{\lambda}\right)^{\zeta'}\right)$	$\frac{\Gamma(p/\zeta'+\zeta/\zeta')}{\Gamma(\zeta/\zeta'')}\lambda^p$	
Weibull, $\underline{x} \ge 0$	$F(x) = 1 - \exp\left(-\left(\frac{x}{\lambda}\right)^{\zeta}\right)$	$\Gamma\left(rac{p}{\zeta}+1 ight)\lambda^p$	
Normal, $\underline{x} \in \mathbb{R}$	$f(x) = \frac{\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma}$		$\begin{cases} \mu_1' = \mu, & p = 1 \\ \sigma^2 & p = 2 \\ 0 & p > 2 \end{cases}$
Half-normal, $\underline{x} \ge 0$	$f(x) = \frac{2}{\lambda\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\lambda^2}\right)$	$\frac{2^{p/2}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right)\lambda^p$	
Extended half-normal (Chi), $\underline{x} \ge 0$	$f(x) = \frac{\sqrt{2}}{\lambda \Gamma(\zeta/2)} \left(\frac{x^2}{2\lambda^2}\right)^{\frac{\zeta}{2} - \frac{1}{2}} \exp\left(-\frac{x^2}{2\lambda^2}\right)$	$2^{p/2} \frac{\Gamma\left(\frac{p+\zeta}{2}\right)}{\Gamma\left(\frac{\zeta}{2}\right)} \lambda^p$	
Lognormal $(\ln \underline{x} \sim (\ln \lambda, \varsigma)), \underline{x} \ge 0$	$f(x) = \frac{\exp\left(-\frac{1}{2\varsigma^2}\left(\ln\left(\frac{x}{\lambda}\right)\right)^2\right)}{\sqrt{2\pi}\varsigma x}$	$e^{rac{p^2 \zeta^2}{2} \lambda^p}$	
Extreme value type I (EV1), $\underline{x} \in \mathbb{R}$	$F(x) = \exp\left(-e^{-\frac{x}{\lambda}}\right)$		$\begin{cases} \gamma \lambda & p = 1 \\ (p-1)! \zeta(p) \lambda^p & p > 1 \end{cases}$

Name, domain	Probability density or distribution function	Moments, μ_p'
Pareto $\underline{x} \ge 0$	$F(x) = 1 - \left(1 + \xi \frac{x}{\lambda}\right)^{-\frac{1}{\xi}}$	$B\left(\frac{1}{\xi}-p,p+1\right)\frac{\lambda^p}{\xi^{p+1}}$
Pareto-Burr-Feller (PBF) $\underline{x} \ge 0$	$F(x) = 1 - \left(1 + \xi \zeta \left(\frac{x}{\lambda}\right)^{\zeta}\right)^{-\frac{1}{\xi\zeta}}$	$B\left(\frac{1}{\xi\zeta} - \frac{p}{\zeta}, \frac{p}{\zeta} + 1\right) \frac{\lambda^p}{\left(\xi\zeta\right)^{\frac{p}{\zeta} + 1}}$
Dagum $\underline{x} \ge 0$	$F(x) = \left(1 + \frac{1}{\xi\zeta} \left(\frac{x}{\lambda}\right)^{-\frac{1}{\xi}}\right)^{-\xi\zeta}$	$(\xi\zeta)^{1-\xi p} \mathbf{B}(1-\xi p,\xi p+\xi\zeta)\lambda^p$
Extreme value type II (EV2) $\underline{x} \ge 0$	$F(x) = \exp\left(-\xi \left(\frac{x}{\lambda}\right)^{-\frac{1}{\xi}}\right)$	$\Gamma(1-p\xi)\left(\frac{\lambda}{\xi}\right)^p$
Half Student $\underline{x} \ge 0$	$f(x) = \frac{2\left(1 + \left(\frac{x}{\lambda}\right)^2\right)^{-\frac{1}{2} - \frac{1}{2\xi}}}{\lambda B\left(\frac{1}{2}, \frac{1}{2\xi}\right)}$	$\frac{\mathrm{B}\left(\frac{1}{2}+\frac{p}{2},\frac{1}{2\xi}-\frac{p}{2}\right)}{\mathrm{B}\left(\frac{1}{2},\frac{1}{2\xi}\right)}\lambda^{p}$
Half extended Student $\underline{x} \ge 0$	$f(x) = \frac{2\left(\left(\frac{x}{\lambda}\right)^2\right)^{\frac{\zeta}{2}-\frac{1}{2}}\left(1+\left(\frac{x}{\lambda}\right)^2\right)^{-\frac{\zeta}{2}-\frac{1}{2\xi}}}{\lambda B\left(\frac{\zeta}{2},\frac{1}{2\xi}\right)}$	$\frac{\mathrm{B}\left(\frac{1}{2\zeta}+\frac{p}{2},\frac{1}{2\xi}-\frac{p}{2}\right)}{\mathrm{B}\left(\frac{1}{2\zeta},\frac{1}{2\xi}\right)}\lambda^{p}$
Generalized beta prime (GBP) $\underline{x} \ge 0$	$f(x) = \frac{\varsigma\left(\frac{x}{\lambda}\right)^{\zeta-1} \left(1 + \left(\frac{x}{\lambda}\right)^{\varsigma}\right)^{-\frac{\zeta}{\varsigma} - \frac{1}{\xi\varsigma}}}{\lambda B\left(\frac{\zeta}{\varsigma}, \frac{1}{\xi\varsigma}\right)}$	$\frac{\mathrm{B}\left(\frac{\zeta}{\varsigma} + \frac{p}{\varsigma}, \frac{1}{\xi\varsigma} - \frac{p}{\varsigma}\right)}{\mathrm{B}\left(\frac{\zeta}{\varsigma}, \frac{1}{\xi\varsigma}\right)} \lambda^{p}$

Table 4. Noncentral moments and cumulants of common distributions with upper-tail index ξ (moments and cumulants exist for $p < 1/\xi$). Here the cumulants do not have simple explicit expressions but can be readily calculated from Equation (9).

- 2. The meaning of constants and standard functions is this: γ : Euler constant; B_p: Bernoulli number of order *p*; $\delta(x)$ the Dirac delta function of *x*; $\Gamma(a)$: gamma function of *a*; B(*a*, *b*): beta function of *a*, *b*.
- 3. Distributions named "half" have their "full" version whose density f(x) and exceedance $\overline{F}(x)$ are obtained by dividing those given in the tables by 2. The "half" version given in the tables corresponds to $\underline{x} \ge 0$, while in the "full" version $\underline{x} \in \mathbb{R}$. The moments μ'_p of the "full" version is: (a) for even p, 0; (b) for odd p, equal to those of half version.
- 4. All other distributions, defined for $\underline{x} \ge 0$ but not named "half", can also be extended to the whole real line by replacing x with |x| and dividing f(x) by 2. Again, the moments μ'_p of this extended version is: (a) for even p, 0; (b) for odd p, equal to those of original version.

2.3 Second order properties

For a stochastic process $\underline{x}(t)$ in continuous time t or \underline{x}_{τ} in discrete time τ , we define the cumulative process $\underline{X}(k) \equiv \underline{X}_{\kappa}$, for continuous time scale $k \coloneqq \kappa D$, where κ denotes discrete time scale, as:

$$\underline{X}(k) \equiv \underline{X}_{\kappa} \coloneqq \underline{x}_{1} + \underline{x}_{2} + \dots + \underline{x}_{\kappa} = \int_{0}^{\kappa D} \underline{x}(t) dt$$
(17)

The time average of the original process \underline{x}_{τ} for discrete time scale κ is

$$\underline{x}_{\tau}^{(\kappa)} \coloneqq \frac{\underline{x}_{(\tau-1)\kappa+1} + \underline{x}_{(\tau-1)\kappa+2} + \dots + \underline{x}_{\tau\kappa}}{\kappa} = \frac{\underline{X}_{\tau\kappa} - \underline{\tilde{X}}_{(\tau-1)\kappa}}{\kappa}$$
(18)

The variability of the time-averaged process is quantified by the variance:

$$\gamma_{\kappa} \coloneqq \operatorname{var}\left[\underline{x}_{\tau}^{(\kappa)}\right] \tag{19}$$

This can be extended to a continuous-time process, for which

$$\gamma(k) \coloneqq \operatorname{var}\left[\frac{\underline{X}(k)}{k}\right], \qquad \gamma_{\kappa} = \gamma(k/D)$$
 (20)

Clearly, this is a function of the time-scale κ which is termed the climacogram of the process, from the Greek climax ($\kappa\lambda i\mu\alpha\xi$, meaning scale) [41].

For sufficiently large κ (theoretically as $\kappa \to \infty$), we may approximate the climacogram as:

$$\gamma(k) \propto k^{2H-2} \tag{21}$$

where *H* is termed the Hurst parameter. The theoretical validity of such (power-type) behaviour of a process was implied by Kolmogorov (1940) [42,43]. The quantity 2H–2 is visualized as the slope of the double logarithmic plot of the climacogram for large time-scales. In a purely random process, H = 1/2, while in most natural processes $1/2 \le H \le 1$, as first observed by Hurst in 1951 [44]. This natural behaviour is known as LRD, (long-term) persistence or Hurst-Kolmogorov (HK) dynamics. A high value of *H* (approaching 1) indicates enhanced presence of patterns, enhanced change and enhanced uncertainty (e.g., in future predictions). A low value of *H* (< 1/2) indicates enhanced fluctuation or antipersistence.

A stochastic process $\underline{x}(t)$ for which the property (21) is valid not only asymptotically but precisely for any scale k, i.e.,

$$\gamma(k) = \lambda \left(\frac{\alpha}{k}\right)^{2-2H}$$
(22)

where α and λ are scale parameters with units of time and $[x^2]$, respectively, is termed the *Hurst-Kolmogorov* (HK) *process* [11].

The HK process is a simple mathematical model offering acceptable approximations for large scales, but it is not physically plausible for small scales because it yields infinite variance of the instantaneous process (as $k \rightarrow 0$) [45]. Therefore, *filtered* versions thereof (FHK) with finite variance at all scales are better options to model natural processes. Here we use two versions of FHK. namely:

• The generalized Cauchy-type FHK (FHK-C) with climacogram:

$$\gamma(k) = \lambda_0 (1 + (k/\alpha)^{2M})^{\frac{H-1}{M}}$$
(23)

• The mixed Cauchy-Dagum-type FHK (FHK-CD) climacogram:

$$\gamma(k) = \lambda_1 \left(1 + \frac{k}{\alpha} \right)^{2H-2} + \lambda_2 \left(1 - \left(1 + \frac{\alpha}{k} \right)^{-2M} \right)$$
(24)

In addition to the Hurst parameter *H*, which characterizes the global scaling behaviour, when $k \to \infty$, the filtered models include a second scaling exponent *M* characterizing the local scaling (or smoothness or fractal behaviour) when $k \to 0$. Furthermore, the FHK-CD model contains two scale parameters of state, λ_1 and λ_2 , instead of the single λ of the FHK-C, offering greater flexibility.

Once the model climacogram is given, all other second-order properties of the process are uniquely determined through simple mathematical expressions. Thus, the autocovariance function in continuous and discrete time, for lags *h* and $\eta = h/D$, respectively, is derived from the climacogram through the relationships [56]:

$$c(h) \coloneqq \operatorname{cov}[x(t), x(t + h)] = \frac{1}{2} \frac{\mathrm{d}^2 h^2 \gamma(|h|)}{\mathrm{d}h^2}$$
(25)

for continuous time and

$$c_{\eta} \coloneqq \operatorname{cov}[x_{\tau}, x_{\tau+\eta}] = \frac{(\eta+1)^2 \gamma_{|\eta+1|} + (\eta-1)^2 \gamma_{|\eta-1|}}{2} - \eta^2 \gamma_{|\eta|}$$
(26)

for discrete time, where cov[] stands for covariance.

Finally, the power spectrum s(w) of the process is the Fourier transform of the autocovariance, so that:

$$s(w) \coloneqq 4 \int_{0}^{\infty} c(h) \cos(2\pi wh) \, \mathrm{d}h \Leftrightarrow c(h) = \int_{0}^{\infty} s(w) \cos(2\pi wh) \, \mathrm{d}w \tag{27}$$

for continuous time and

$$s_{\rm d}(\omega) = 2c_0 + 4\sum_{\eta=1}^{\infty} c_\eta \cos(2\pi\eta\omega) \Leftrightarrow c_\eta = \int_0^{1/2} s_{\rm d}(\omega) \cos(2\pi\omega\eta) \,\mathrm{d}\omega \tag{28}$$

for discrete time.

2.4 Stochastic simulation

To simulate the discrete-time stochastic process \underline{x}_{τ} with any autocovariance function c_{η} we can use the generalized moving average scheme [10,11,26]:

$$\underline{x}_{\tau} = \sum_{j=-J}^{J} a_j \underline{v}_{\tau-j} \tag{29}$$

where a_j are weights to be calculated from the autocovariance function, \underline{v}_j is white noise averaged in discrete-time (in the general case assumed non-Gaussian), also known as *innovation process*, and *J* is theoretically infinite, so that in all theoretical calculations we assume $J = \infty$, while in the generation phase *J* is a large integer chosen so that the resulting truncation error be negligible.

Here we stress that the above scheme is just the contrary to the common schemes of the Time Series School. Specifically, (a) we use a purely moving average scheme without any autoregressive term and (b) we do not relate our scheme with observations, as the observations have been already used in the model fitting phase, which is totally isolated from the generation scheme. Specifically, the fitting constitutes of a choice of an appropriate climacogram expression such as (23) or (24) and the estimation of its parameters, as well as the choice of a distribution function such as those contained in Table 2- Table 4 and the estimation of its parameters. This tactic assures modelling parsimony. More details on the fitting procedure, which is not covered here, can be found in [31]. Here we only stress the methodological suggestion that we never estimate from data classical moments and cumulants of order greater than 2, because these are *unknowable* from data [46]. While the methodology that will follow heavily depends on high-order moments and cumulants, it is stressed that these are determined by theoretical calculations and never from the data.

Assuming unit variance of the white noise \underline{v}_j , writing Equation (29) for $\underline{x}_{\tau+\eta}$, multiplying it by (29) and taking expected values we find the convolution expression for $J = \infty$:

$$c_{\eta} = \sum_{l=-\infty}^{\infty} a_l a_{\eta+l} \tag{30}$$

We need to find the sequence of a_{η} , $\eta = \cdots$, -1,0,1, ..., so that (30) holds true. The following generic solution of the generating scheme, giving the coefficients a_{η} , has been proposed by Koutsoyiannis [26]:

$$a_{\eta} = \int_{-1/2}^{1/2} e^{2\pi i (\vartheta(\omega) - \eta \omega)} A^{R}(\omega) d\omega$$
(31)

where $i \coloneqq \sqrt{-1}$, $\vartheta(\omega)$ is any (arbitrary) odd real function (meaning $\vartheta(-\omega) = -\vartheta(\omega)$) and

$$A^{\mathrm{R}}(\omega) \coloneqq \sqrt{2s_{\mathrm{d}}(\omega)} \tag{32}$$

As proved by Koutsoyiannis [26], the sequence of a_{η} :

- 1. consists of real numbers, despite the expression in (31) involving complex numbers;
- 2. satisfies precisely equation (30); and
- 3. is easy and fast to calculate using the fast Fourier transform (FFT).

This theoretical result is readily converted into a numerical algorithm, which consists of the following steps [26]:

- 1. From the continuous-time stochastic model, expressed through its climacogram $\gamma(k)$, we calculate its autocovariance function in discrete time (assuming time step *D*) by Equation (26). (This step is obviously omitted if the model is already expressed in discrete time through its autocovariance function.)
- 2. We choose an appropriate number of coefficients *J* that is a power of 2 and perform inverse FFT (using common software) to calculate the discrete-time power spectrum and the frequency function $A^{R}(\omega)$ for an array of $\omega_{i} = j w_{1}, j = 0, 1, ..., J, w_{1} \coloneqq 1/JD$:

$$s_{\rm d}(\omega_j) = 2c_0 + 4\sum_{\eta=1}^J c_\eta \cos(2\pi\eta\omega_j), \qquad A^{\rm R}(\omega_j) = \sqrt{2s_{\rm d}(\omega_j)}$$
(33)

3. We choose $\vartheta(\omega)$ (see below) and we form the arrays (vectors) A^{R} and A^{I} , both of size 2*J* indexed as 0, ..., 2*J* – 1, with the superscripts R and I standing for the real and imaginary part of a vector of complex numbers, respectively:

$$[\mathbf{A}^{\mathrm{R}}]_{j} = \begin{cases} A^{\mathrm{R}}(\omega_{j})\cos\left(2\pi\vartheta(\omega_{j})\right)/2, & j = 0, \dots, J\\ [\mathbf{A}^{\mathrm{R}}]_{2J-j}, & j = J+1, \dots, 2J-1 \end{cases}$$
(34)

$$[\mathbf{A}^{\mathrm{I}}]_{j} = \begin{cases} -A^{\mathrm{R}}(\omega_{j})\sin(2\pi\vartheta(\omega_{j}))/2, & j = 0, \dots, J-1\\ 0 & j = J\\ -[\mathbf{A}^{\mathrm{I}}]_{2J-j}. & j = J+1, \dots, 2J-1 \end{cases}$$
(35)

4. We perform FFT on the vector A^{R} + i A^{I} (using common software), and get the real part of the result for j = 0, ..., J, which is precisely the sequence of a_{η} .

We note that by choosing *J* as a power of 2, the vectors A^{R} and A^{I} will have size 2*J* which is also a power of 2, thus achieving maximum speed in the FFT calculations. (More details are contained in a supplementary file in [26], which includes numerical examples along with the simple code needed to do these calculations on a spreadsheet).

Remarkably, Equation (31) gives not a single solution, but a family of infinitely many solutions, all of which preserve exactly the second-order characteristics of the process. A particular solution is characterized by the chosen function $\vartheta(\omega)$. Even assuming $\vartheta(\omega) = \vartheta_0 \operatorname{sign} \omega$ with constant ϑ_0 , again there are infinitely many solutions. Furthermore, if the sequence of $\vartheta(\omega_j)$ is constructed at random (e.g., as a sequence of random numbers in the interval [0,1/4]), again equation (30) will be satisfied and the resulting a_{η} can be directly used in generation. The availability of infinitely many solutions enables preservation of additional statistics, such as those related to time asymmetry [26].

The special case $\vartheta(\omega) = 0$ gives a symmetric solution with respect to positive and negative η :

$$A^{\rm S}(\omega) \equiv A^{\rm R}(\omega) = \sqrt{2s_{\rm d}(\omega)}, \qquad a_j^{\rm S} = \int_0^{1/2} \sqrt{2s_{\rm d}(\omega)} \cos(2\pi j\omega) \,\mathrm{d}\omega = a_{-j}^{\rm S}$$
(36)

where the superscript S stands for symmetric. This has been known as the *symmetric moving average* (SMA) scheme [10]. All other solutions denote *asymmetric moving average* (AMA) schemes.

In addition, we always have several options related to the distribution of the white noise \underline{v}_{τ} , which in general is not Gaussian, thus enabling preservation of moments and cumulants of any order. Specifically, the *p*th cumulants of the processes \underline{x}_{τ} and \underline{v}_{τ} , by virtue of (13), are related by:

$$\kappa_p = \sum_{j=-J}^{J} a_j^p \, \kappa_p^{(v)} \tag{37}$$

where κ_p and $\kappa_p^{(v)}$ are *p*th cumulants of \underline{x}_{τ} and \underline{v}_{τ} , respectively. Solving for $\kappa_p^{(v)}$ we find:

$$\kappa_p^{(\nu)} = \frac{\kappa_p}{\sum_{l=-J}^J a_j^p} \tag{38}$$

Given the so-calculated $\kappa_p^{(v)}$ for any order p, the distribution function of the white noise is fully determined. However, in practice we can only approximate this distribution function up to an order p_{max} . A convenient way to make the approximation is to choose a number L of elementary distribution functions from Table 2- Table 4, thus defining the white-noise processes \underline{w}_l , l = 1, ..., L, and obtain the approximation \underline{v}_r' of \underline{v}_r as a linear combination of \underline{w}_l with weights a'_l , i.e.:

$$\underline{v}_{x}' = \sum_{l=1}^{L} a_{l}' \underline{w}_{l} \tag{39}$$

The cumulants $\kappa_p^{(w_l)}$ of \underline{w}_l are then determined from Table 2- Table 4 and those of \underline{v}'_{τ} , by virtue of (13), are:

$$\kappa_p^{(v')} = \sum_{l=-J}^{J} {a'_j}^p \, \kappa_p^{(w_l)} \tag{40}$$

The goodness of the approximation up to order p_{max} is given by an error expression such as:

$$e_{1} \coloneqq \sum_{p=2}^{p_{\max}} \left(\left(\kappa_{p}^{(v')} \right)^{\frac{1}{p}} - \left(\kappa_{p}^{(v)} \right)^{\frac{1}{p}} \right)^{2}, \qquad e_{2} \coloneqq \sum_{p=2}^{p_{\max}} \left(\frac{1}{p} \ln \left(\frac{\kappa_{p}^{(v')}}{(\kappa_{p}^{(v)})} \right) \right)^{2}$$
(41)

where the second form (e_2) is more appropriate if all cumulants are positive and increasing fast. In order for the above equations to work in all cases, even when κ_p is negative and p is even, the quantity $(\kappa_p)^{1/p}$ is meant to denote the quantity $sign(\kappa_p) |\kappa_p|^{1/p}$; this convention is followed throughout the entire paper. By minimizing either e_1 or e_2 using a common solver, we simultaneously find the series of weights a'_l and the parameters of the marginal distribution of each of \underline{w}_l . Further details will be given in the applications of section 3, where it will be also seen that, for a sufficient approximation, the number of constituent distributions L of \underline{w}_l is small, usually 1 or 2.

It is stressed that, in each of the above error expressions, we have intentionally excluded the error of the cumulants of order 1, i.e. the mean values. Therefore, we expect that with this procedure the mean will not be preserved. However, this can be easily tackled by adding a constant c to \underline{v}'_{r} . Apparently, the required shift should be

$$c = \kappa_1^{(v)} - \kappa_1^{(v')} \tag{42}$$

Based on the above approximation, the generation process will produce the stochastic process

$$\underline{x}'_{\underline{x}} \coloneqq \sum_{j=-1}^{J} a_j \underline{v}'_{\tau-j} \tag{43}$$

where, if the approximation is satisfactory, we reasonably expect that the statistical properties of \underline{x}'_{τ} will be equal to those of \underline{x}_{τ} . This proves to be always the case if the domain of the stochastic variable \underline{x}_{τ} is unbounded in both directions (i.e., $\underline{x}_{\tau} \in \mathbb{R}$) but some additional manipulation (post processing) may be needed if the domain of \underline{x}_{τ} is not the entire real line, or if the distribution function of \underline{x}_{τ} has discontinuities, as will be illustrated in the applications of the next section.

3. Applications and results

We illustrate the methods by five applications for bounded \underline{x}_{τ} as this case is more demanding (the unbounded case is much easier). Three applications are synthetic mathematical examples used as benchmarks, namely the exponential distribution, which is bounded from below, and the uniform distribution, which is bounded from both below and above. The next two are real-world applications dealing with one of the most challenging natural processes, namely the precipitation process which is bounded from below, highly intermittent and with heavy distribution tail. The latter two applications refer to two different time scales, fine (hourly) and coarse (annual). In the synthetic example with the exponential distribution and in the two real-world applications, the stochastic processes are persistent with a large Hurst parameter, ranging from 0.80 to 0.92. In the synthetic examples of the uniform distribution we use both a persistent and an antipersistent process, with Hurst parameters 0.70 and 0.20, respectively.

3.1. Simulating a persistent process with exponential distribution

For a process with exponential distribution, which is a subcase of the gamma distribution, there exist generation algorithms for the case of short-range (Markov) dependence (e.g. [47]). As already mentioned, a more general algorithm for generation of any type of marginal distribution has recently been proposed by Lombardo et al. [28] but again under the condition of Markov dependence. However, the method proposed here can generate such a process irrespective of the type of the dependence, whether SRD or LRD.

For illustration we assume an FHK-C model (Equation (23)) with parameters H = 0.8, M = 0.5, $\alpha = 1$, $\lambda_0 = 1.32$, so that $\gamma_1 = 1$. The FHK-C climacogram is shown in Figure 1(b), marked as "theoretical", while the resulting autocorrelation function is shown in Figure 1(c). As in the exponential distribution (from Table 3), $\mu = \sqrt{\gamma_1} = 1$, the cumulants of the process \underline{x}_r are $\kappa_p = (p - 1)!$. These are depicted in Figure 1(a), along with the cumulants of \underline{v}_r determined from Equation (38), where, to avoid big numbers, the quantities $\kappa_p^{1/q}$ are plotted. The coefficients a_j , needed to evaluate $\kappa_p^{(v)}$ in Equation (38), are determined from the SMA (symmetric) generation scheme (Equation (36)) with J = 1024.

Coming to the approximation \underline{v}'_{t} of \underline{v}_{t} , we use two constituents \underline{w}_{l} with gamma distributions and allow a discontinuity P_{1} at w = 0 in each of them. Assuming unit variance in each of them, from the equations of Table 3 we have $\zeta \lambda^{2} = 1$, so that the continuous part of the distribution is fully determined by the shape parameter ζ . Hence the approximation \underline{v}'_{t} , according to Equation (39), is determined by the parameters ζ_{1} , ζ_{2} , P_{1} , P_{2} , a'_{1} , a'_{2} , which are calculated by minimizing e_{2} in Equation (41), assuming $p_{\text{max}} = 10$. The resulting values of the parameters are $\zeta_{1} = 1.255$, $\zeta_{2} = 30$, $P_{1} = 0.298$, $P_{2} = 1$, $a'_{1} = 1.333$, $a'_{2} = -0.0655$, while the required shift of Equation (42) is negligible (c = 0). The cumulants of

 \underline{v}_{τ}' are also plotted in Figure 1(a), where it can be seen that they are indistinguishable from those of \underline{v}_{τ} and thus the achieved approximation is very good.

The generation of values of \underline{v}'_{τ} is quite easy using a random number generator for the gamma distribution. From a series of random numbers v'_{τ} , a total of $n = 10\,000$ values of \underline{x}_{τ} are then determined from Equation (29). A small number (6.6%) of them are small negative values. To remedy this problem, we reflect these values about zero, or, in other words, replace x_{τ} with $-x_{\tau}$. Theoretically, this remedy will have a distorting effect in the multivariate distribution of \underline{x}_{τ} , but in fact this effect turns out to be negligible.



Figure 1. Graphical depiction of the results of the simulation application for a synthetic example of a persistent FHK process with exponential distribution: (a) cumulants; (b) climacogram; (c) autocorrelogram; (d) marginal distribution.

Comparison of the theoretical statistical characteristics of the distribution of \underline{x}_r to the empirical ones of the generated sample are shown in the panels of Figure 1. In the empirical climacogram (Figure 1(b)), the plotted points correspond to unbiased estimates of variance; this is achieved by adding the quantity $\gamma(n) = 0.0331$ to the classical statistical estimates, as explained in [31]. The empirical climacogram agrees well with the theoretical one. The empirical autocorrelation is shown in Figure 1(c). Here the bias correction was applied using an approximate method from [48], according to which the unbiased estimate is the weighted sum of the classical autocorrelation estimate and the number 1, with the weight of the latter being equal to 1/n', where $n' \coloneqq \gamma(1)/\gamma(n)$ is the so-called

equivalent sample size of any process, and differs substantially from *n* if the process is persistent [49]. (We note that a precisely unbiased estimate of autocovariance has been provided by [50] but this is more laborious). Finally, Figure 1(d) shows a comparison of the theoretical and empirical marginal distribution of \underline{x}_{τ} . The empirical distribution of each value of the generated time series, arranged in ascending order so that $x_{(i:n)}$ be the *i*th smallest value of the series of *n* values, was estimated on the basis of unbiasedness of the logarithm of return period $T_{(i:n)}$. As shown in [31], this estimate is

$$\frac{T_{(i:n)}}{D} = \frac{n + e^{1-\gamma} - 1}{n - i + e^{-\gamma}} = \frac{n + 0.526}{n - i + 0.561}$$
(44)

Again, the agreement between theoretical and the empirical distributions is very good.

3.2. Simulating a persistent process with uniform distribution

The simulation of a persistent process with uniform distribution is more demanding because of the double boundedness and the sharp discontinuities of the density function at the bounds, while linear generation procedures tend to generate unbounded processes with smooth density. On the other hand, the double boundedness offers an option of approximation with a process \underline{v}'_{τ} that takes on a finite number of values. In other words, we assume that the stochastic variable \underline{v}'_{τ} is discrete, taking on a finite number of values v'_{i} with probabilities P_{i} , as illustrated in Figure 2. The details of this approximation will be explained in a while. Despite v'_{τ} assumed discrete, thanks to the fact that the generation of \underline{x}_{τ} via Equation (29) involves a linear combination of very many variables v'_{τ} , the variable \underline{x}_{τ} will in effect be continuous.



Figure 2. Probability mass function of the discretized white noise used in the simulation application for a synthetic example of a persistent FHK process with uniform distribution.

As in the previous case, for illustration we assume an FHK-C model (Equation (23)) with $\gamma_1 = 1$. We note that the fourth cumulant of this uniform distribution, which in this case equals the coefficient of kurtosis, is $\kappa_4 = -1.2$. The fourth cumulant of \underline{v}_r ($\kappa_4^{(v)}$) should necessarily be lower than that ($\kappa_4^{(v)} < -1.2$) for a persistent process. On the other hand, it is known than the kurtosis of any distribution cannot be lower than -2. Therefore, the margin for having a positively autocorrelated process \underline{x}_r with uniform distribution is rather small. An FHK-C model with parameters H = M = 0.7, $\alpha = 1$, $\lambda_0 = 1.346$ (so that $\gamma_1 = 1$) yields a feasible $\kappa_4^{(v)} = -1.76$, while the case H = M = 0.75, would yield an infeasible $\kappa_4^{(v)} = -2.02$. The FHK-C climacogram for the feasible parameter set (H = M = 0.7) is shown in Figure 3(b), marked as "theoretical", while the resulting autocorrelation

function is shown in Figure 3(c). In order for the uniform distribution to have variance $\gamma_1 = 1$, its upper bound should be $b = \sqrt{12} = 3.464$, with lower bound a = 0. The cumulants of the process \underline{x}_r , determined from Table 2 and Equation (9), are shown in Figure 3(a), along with the cumulants of \underline{v}_r determined from Equation (38) (for the convention used for $\kappa_p^{1/q}$ for negative quantities and p even, see the note of section 2.4 below Equation (41)). The coefficients a_j , needed to evaluate $\kappa_p^{(v)}$ in Equation (38), are determined from the SMA (symmetric) generation scheme (Equation (36)) with J = 1024.



Figure 3. Graphical depiction of the results of the simulation application for a synthetic example of a persistent FHK process with uniform distribution: (a) cumulants; (b) climacogram; (c) autocorrelogram; (d) marginal distribution.

Comparisons of the theoretical statistical characteristics of the distribution of \underline{x}_{τ} to the empirical ones of the generated sample are shown in the panels of Figure 3, which are similar as those in Figure 1. A difference is that in panel (d), instead of estimating the return period of each $x_{(i:n)}$ (the *i*th smallest value of the series of *n* values), we give the non-exceedance probability F(x), estimated on the basis of its unbiasedness. In this case, the unbiased estimate is [31]:

$$F(x_{(i:n)}) = \frac{i}{n+1} \tag{45}$$

In all panels, the agreement between theoretical and the empirical characteristics is very good.

The approximation \underline{v}'_{r} of \underline{v}_{r} is done through the discretization of the former described above. Twenty equidistant v'_{i} with probabilities P_{i} are assumed, where $v'_{i} = i/b$, i = 1, ..., 20. The distribution of v'_{i} was assumed symmetric, i.e., $P_{i} = P_{21-i}$, so that the unknown parameters to be optimized are ten, namely, $P_{1}, ..., P_{10}$. These are calculated by minimizing e_{1} in Equation (41), assuming $p_{max} = 10$. The resulting values are shown graphically in Figure 2. It is remarkable that the distribution of \underline{v}'_{r} is far from uniform, despite the fact that the cumulants of \underline{v}'_{r} , as seen in Figure 3(a), are not very different from those of \underline{x}_{τ} , which has uniform distribution. The cumulants of \underline{v}'_{r} , also plotted in Figure 1(a), are indistinguishable from those of \underline{v}_{r} , and thus the achieved approximation is very good. An exception is seen in the first cumulants of \underline{v}'_{r} and \underline{v}_{r} , which are quite different, and thus the required shift of Equation (42) is not negligible, namely c = -1.503.

The generation phase is quite easy, as values of \underline{v}'_{τ} are readily generated by inverse transform sampling, given the staircase-like distribution function of a discrete stochastic variable. A total $n = 10\ 000$ values of \underline{x}_{τ} are then generated from Equation (29). A small number (~2%) of them are either small negative values or somewhat greater than b. As in the previous case, we reflect the negative values about zero, replacing x_{τ} with $-x_{\tau}$. Likewise, we reflect the very high values about b, replacing x_{τ} with $2b - x_{\tau}$.

3.3. Simulating as antipersistent process with uniform distribution

For further illustration we examine the same uniform distribution as above but for an antipersistent process (with H < 1/2). Actually, this case is easier as the changes in kurtosis is smaller than in the previous case, and thus, feasibility of the solution is assured.

Again, an FHK-C model was assumed, now with parameters H = 0.2, M = 0.8, $\alpha = 1$, $\lambda_0 = 2$ (so that $\gamma_1 = 1$, while $\kappa_4^{(v)} = -1.265$). All other choices are the same as in the previous application (e.g. upper bound $b = \sqrt{12} = 3.464$, etc.) The approximation \underline{v}'_r of \underline{v}_r through discretization is depicted in Figure 4. Again, this differs substantially from the uniform distribution, even though the cumulants of \underline{v}'_r , as seen in Figure 5(a), are virtually indistinguishable from those of \underline{x}_r and \underline{v}_r . Yet there is a substantial difference in the first cumulants of \underline{v}'_r and \underline{v}_r , so that the required shift of Equation (42) is large, c = 13.675.



Figure 4. Probability mass function of the discretized white noise used in the simulation application for a synthetic example of an antipersistent FHK process with uniform distribution.

Comparisons of the theoretical statistical characteristics of the distribution of \underline{x}_{τ} to the empirical ones of the generated sample are shown in the panels of Figure 5. In all panels the agreement between theoretical and the empirical characteristics is very good.



Figure 5. Graphical depiction of the results of the simulation application for a synthetic example of an antipersistent FHK process with uniform distribution: (a) cumulants; (b) climacogram; (c) autocorrelogram; (d) marginal distribution. Notice in panel (a) that the first cumulant of \underline{v} is out of the graph area as it is very large ($\kappa_1^{(v)} = 15.49$).

3.4. Simulating the precepitation process at the hourly time scale

Here we use a recently developed [31] full stochastic model of the precipitation process at any time scale *k*. This model gives directly the ombrian relationships (else known as intensity-duration-frequency curves) but it also provides any stochastic characteristic of the precipitation process that is required for stochastic simulation. Furthermore in [31] this model has been applied to construct the ombrian curves by fitting the model in some locations, but the model was not used for stochastic simulation. Among the locations studied in [31], here we provide a stochastic simulation for rainfall in Bologna, using the parameter valued fitted there. The application in this subsection is for the hourly scale, while an additional application for the annual scale is given in the next subsection.

The model is based on the following assumptions, which are mathematically consistent (with one exception as detailed below):

1. Pareto distribution with discontinuity at the origin for small time scales (Table 5, Equation (46), left). The tail index ξ is constant for all time scales k, while the probability wet, $P_1^{(k)}$, and the state scale parameter, $\lambda(k)$, are functions of the time scale k.

- 2. Continuous PBF distribution, possibly with discontinuity at zero, for large time scales (Table 5, Equation (46), right). In this case, a new parameter $\zeta(k)$ is introduced, which is again a function of time scale. The Pareto distribution is a special case of the PBF for $\zeta(k) = 1$. In contrast to the Pareto distribution, whose density is a consistently decreasing function of *x*, the PBF tends to be bell-shaped for increasing $\zeta(k)$, a property consistent with empirical observation and reason.
- 3. Constant mean μ of the time-averaged process.
- 4. Climacogram of type FHK-CD (Equation (24)), where to reduce the number of parameters it is assumed that M = 1 H, thus getting Equation (48) in Table 5. By inspection of Equation (48), it is seen that, as $k \to \infty$, $\gamma(k) \to 0$, which makes the process ergodic; for k = 0, $\gamma(0) = \gamma_0 = \lambda_1 + \lambda_2$, which is finite, as required for physical consistency.
- 5. Probability wet and dry, $P_1^{(k)} = 1 P_0^{(k)}$, varying with time scale according to Equation (49) in Table 5. It is clarified that two different expressions are used for the small and the large scales, where the transition time scale from the Pareto to the PBF distribution is denoted as k^* . In the Pareto case, $P_1^{(k)}$ can be determined directly from the climacogram and the mean (left column of Equation (49) in Table 5). For the PBF case an additional equation is required, which has been derived based on maximum entropy considerations [51] and involves an additional parameter θ ($0 \le \theta \le 1$). Continuity of the transition demands that $\zeta(k^*) = 1$.

Quantity and symbol	Small scales, $k \leq k^*$ (Pareto)	Large scales, $k \ge k^*$ (PBF)	Eqn. no.
Distribution function, $F^{(k)}(x)$	$1 - P_1^{(k)} \left(1 + \xi \frac{x}{\lambda(k)} \right)^{-1/\xi}$	$1 - P_1^{(k)} \left(1 + \xi' \zeta(k) \left(\frac{x}{\lambda(k)} \right)^{\zeta(k)} \right)^{-\frac{1}{\xi' \zeta(k)}}$	(46)
Mean, $E[\underline{x}^{(k)}]$		μ	(47)
Climacogram, $\gamma(k)$	$\lambda_1 \left(1 + \frac{k}{\alpha}\right)^{2H}$	$^{-2} + \lambda_2 \left(1 - \left(1 + \frac{\alpha}{k} \right)^{2H-2} \right)$	(48)
Probability wet, $P_1^{(k)}$	$\frac{1-\xi}{1/2-\xi}\frac{\mu^2}{\gamma(k)+\mu^2}$	$1 - \left(1 - P_1^{(k^*)}\right)^{(k/k^*)^{\theta}}$, $(0 \le \theta \le 1)$	(49)
Lower tail index (inverse), $\frac{1}{\zeta(k)}$	1	$\sqrt{(1-2\xi) \left(P_1^{(k)}(\gamma(k)/\mu^2 + 1) - 1 \right)}$	(50)
Upper tail index, ξ	ξ	$\xi' = \frac{\xi}{\zeta(k)}$	
Scale parameter (inverse), $\frac{1}{\lambda(k)}$	$\frac{P_1^{(k)}}{\mu(1-\xi)}$	$\frac{P_1^{(k)}}{\mu} \left(1 + \frac{1}{(1-\xi)(\zeta(k))^2} - \frac{1}{(\zeta(k))^{\sqrt{2}}} \right)$	(51)
Quantile, <i>x</i>	$\lambda(k) \frac{\left(P_1^{(k)} T/k\right)^{\xi} - 1}{\xi}$	$\lambda(k) \left(\frac{\left(P_1^{(k)} T/k \right)^{\xi} - 1}{\xi} \right)^{\frac{1}{\zeta(k)}}$	(52)

Table 5 Mathematical relationships of the ombrian model. The ombrian curves per se are given in the last row.

Both the decreasing (Pareto) and the bell-shaped (PBF) types of probability densities are consistent with natural behaviours for small and large time scales, respectively. It is noted though that the tail index of the PBF distribution in the form in Table 4, is not ξ but $\xi' = \xi/\zeta(k)$ and tends to zero as $k \to \infty$. This violates a requirement of a constant tail index, which is theoretically justified in [31], but this violation happens only for large time scales. The alternative to keep a constant tail index ξ , is not an option because it would result in a finite variance as $k \to \infty$ (with a coefficient of variation $\xi/\sqrt{1-2\xi}$), i.e., in a nonergodic process.

What it remains to complete the model is to determine the functions $\lambda(k)$ and $\zeta(k)$ from the mean μ and the climacogram $\gamma(k)$. This has been done in [31] and the results are shown in Table 5. The final relationships rely on the mean μ , the climacogram $\gamma(k)$, the probability wet $P_1^{(k)}$ and the tail index ξ of the distribution function of rainfall intensity. For precipitation in Bologna the following model parameters have been estimated in [31], while the transition time scale was set $k^* = 96$ h:

- Mean intensity, $\mu = 0.0823$ mm/h;
- Intensity scale parameters, $\lambda_1 = 0.00110 \text{ mm}^2/\text{h}^2$, $\lambda_2 = 1.43 \text{ mm}^2/\text{h}^2$;
- Time scale parameter, α = 8.74 h;
- Hurst parameter, H = 0.92; fractal (smoothness) parameter, M = 1 H = 0.08;
- Exponent of the expression of probability dry/wet, θ = 0.787;
- Upper tail index, $\xi = 0.121$.

For the hourly time scale, the resulting distribution is Pareto (Table 4, Table 5) with a discontinuity at zero, $P_0 \coloneqq P\{x = 0\} = 1 - P_1$ and parameters $\xi = 0.121, \lambda = 2.046 \text{ mm/h}, P_1 = 0.0354$. The FHK-CD climacogram is shown in Figure 6(b) (marked as "theoretical"), while the resulting autocorrelation function is shown in Figure 6(c). The cumulants of the process \underline{x}_r are shown Figure 6(a), along with the cumulants of \underline{v}_r determined from Equation (38). The coefficients a_j , needed to evaluate $\kappa_p^{(v)}$ in Equation (38), are determined from an AMA (asymmetric) generation scheme (Equation (31)) with J = 1024 and phases ϑ generated randomly (this contributes to a realistic shape of generated rainfall events).

For the approximation \underline{v}'_{τ} of \underline{v}_{τ} , we use a single Pareto distribution and allow a discontinuity P_1 at v' = 0. For mathematical consistency, the tail index of \underline{v}'_{τ} should necessarily be $\xi = 0.121$, so that the moments of order beyond $1/\xi = 8.2$ be infinite as is the case with the moments of \underline{x}_{τ} . The other parameters of the Pareto distribution of \underline{v}'_{τ} are calculated by minimizing e_2 in Equation (41), setting $p_{\text{max}} = 8$, and are found to be $\lambda^{(v')} = 3.681, P_1^{(v')} = 0.0171$, while the required shift of Equation (42) is negligible (c = 0). The cumulants of \underline{v}'_{τ} are also plotted in Figure 6(a), where it can be seen that they are indistinguishable from those of \underline{v}_{τ} and thus the achieved approximation is very good.

Because of the very small value of $P_1^{(v')}$, a very large number of \underline{v}'_{τ} (98.3%) will be zero. The nonzero values will determine the locations of rainfall events, i.e., sequences of non-zero x_{τ} . It is not reasonable to make these locations purely random and for this reason we devised the following procedure. A first model run is done with $P_1^{(v')} = 1$ (no zeros). Subsequently, we find a threshold c_0 so that the fraction of values x_{τ} that are greater than c_0 equal $P_1^{(v')}$. In a second model run we set $v'_{\tau} = 0$ at those τ where in the first run $x_{\tau} < c_0$. For the remaining τ we generate v'_{τ} from the continuous part of \underline{v}'_{τ} . This procedure allows clustering of the precipitation events, as typically happens in reality.

The values x_{τ} in the second run, will unavoidably be nonzero, because the generating Equation (29) involves a linear combination of very many v'_{τ} and this can hardly result in zero values. Therefore, post processing of the generated time series is required, in order to reinstate the required number of zeros. This constitutes of replacing x_{τ} by x'_{τ} , determined as:

$$x'_{\tau} = \begin{cases} 0 & x_{\tau} < c_1 \\ l(x_{\tau} - c_1)^m & x_{\tau} \ge c_1 \end{cases}$$
(53)

where c_1 , l and m are the parameters of the post-processing phase. These are determined by minimizing the total error (in effect making it zero) in preserving the probability wet, and the first and second cumulants of the distribution. In our application, the post-processing parameters have been found to be $c_0 = 3.18 \text{ mm/h}$, $c_1 = 1.15 \text{ mm/h}$, l = 1.877, m = 0.832. Comparison of the theoretical statistical characteristics of the distribution of \underline{x}_{τ} to the empirical ones of the generated sample, both before and after post processing, are shown in the panels of Figure 6. The empirical climacogram is shown in Figure 6(b). Before post processing there is a marked difference of the empirical climacogram from the theoretical. This does not indicate a weakness of the algorithm. It just reflects the fact that, with a Hurst parameter as high as H = 0.92, there is high uncertainty and variability, while a sample of $n = 10\ 000$ is too short to eliminate this uncertainty; note that the equivalent sample size (which indicates the sampling variability) in this case is $n' \coloneqq \gamma(1)/\gamma(n) \approx 7$ instead of $n = 10\ 000$. Interestingly, the post processing substantially decreases the difference from the theoretical curve. The improvement due to post processing is spectacular in panel (d), which shows a comparison of the theoretical and empirical marginal distribution of \underline{x}_{τ} . Before post processing, even though the cumulants are preserved, the initially generated small values are problematic as no zero values are generated. This is fully remedied by the post processing technique. Finally, panel (c) shows that the autocorrelations are well preserved both before and after post processing.



Figure 6. Graphical depiction of the results of the simulation application for a real-world case study for the precipitation process in Bologna at the hourly time scale, modelled as a persistent FHK process with Pareto distribution with discontinuity at zero: (a) cumulants; (b) climacogram; (c) autocorrelogram; (d) marginal distribution.

Further information on the form of the generated time series is provided in Figure 7, this time showing not statistical characteristics but the time series per se. The plot, covering a period of 2000 h (83 d; panel (a)) with a focus on the first 200 h (~8 d; panel (b)), indicates that the time series resemble the form of natural rainfall events.



Figure 7. Plots of generated time series of precipitation in Bologna at hourly time scale: (a) for a period of 2000 h (83 d); (b) focus on the first 200 h (~8 d).

3.5. Simulating the precepitation process at the annual time scale

The same precipitation model as in the previous subsection was used for generation at the annual scale. Now the distribution is no longer Pareto but PBF, whose treatment is more laborious. On the other hand, the probability dry at the annual scale is zero, and thus the distribution is continuous. This makes the generation simpler as no post processing is required.

While at the hourly scale all cumulants are positive tending fast to infinity (Figure 6(a)), at the annual scale some of the cumulants (most notably the fourth) are negative (Figure 8(a)). According to the model, again the cumulants tend to infinity but for much higher p (>33) as now $\xi' = 0.030$. The other parameters of the PBF distribution are $\zeta^{(v)} = 4.00$ and $\lambda^{(v)} = 0.089$ mm/h. The approximation \underline{v}'_{t} of \underline{v}_{t} is made by another PBF distribution with slightly different parameters, $\zeta^{(v')} = 4.01$ and $\lambda^{(v')} = 0.098$ mm/h. As seen in Figure 8(a), the achieved approximation is good, except for a substantial difference in the first cumulants of \underline{v}'_{t} and \underline{v}_{r} , so that the required shift of Equation (42) is not negligible, c = 0.0871 mm/h.

Comparisons of the theoretical statistical characteristics of the distribution of \underline{x}_{τ} to the empirical ones of the generated sample are shown in the panels of Figure 8. In all panels the agreement between theoretical and empirical characteristics is very good.



Figure 8. Graphical depiction of the results of the simulation application for a real-world case study for the precipitation process in Bologna at the annual time scale, modelled as a persistent FHK process with PBF distribution: **(a)** cumulants; **(b)** climacogram; **(c)** autocorrelogram; **(d)** marginal distribution.

4. Discussion and conclusions

Stochastic simulation of complex processes necessarily relies on approximations of distribution functions. Typically, these approximations are made with reference to the normal distribution, e.g. the Gram-Charlier series, the Edgeworth approximation, etc. [35-37,52]. These, however, are not good for simulation as no generic random number generation algorithms are available for such type of approximations. They can also be too complicated. Here we provide more general and more powerful approximations of distribution functions on the basis of cumulants. These are quite flexible and can have several forms, such as (a) the sum of a few (e.g. two or even just one) stochastic variables with typical distributions of an appropriate type (such as those contained in Table 2-Table 4), (b) the occasional involvement of discontinuities in constituent distributions (usually at their lower bounds) and (c) the discretization of the stochastic variable, in the case that its domain is bounded from both above and below. As random number generation algorithms are readily available for these typical distributions, the proposed approximation is useful in stochastic simulation.

The approximation of a distribution via cumulants turns out to provide very powerful means for stochastic simulation of processes of any type, with short- and long-range dependence. The combination of this approximation with the asymmetric (AMA) or symmetric (SMA) moving average generation schemes can tackle really demanding simulation processes. The genuine stochastic simulation approach that is studied, which does not perform transformations of the stochastic variables involved, is useful, convenient and powerful. This is particularly the case for problems where time directionality is important; it is reminded that a Gaussian process, even when (back) transformed to non-Gaussian by any nonlinear transformation, cannot provide a process with time asymmetry.

The case studies conducted confirm the excellent performance of the method for a variety of demanding problems and a variety of distributions and time scales. In particular, the long-range dependence, however high, as well as the antipersistence, do not entail any difficulty in applying the method. In contrast, some characteristics of the marginal distribution, such as single or double boundedness, and especially the possible intermittence, may cause difficulties. For this reason, all case studies conducted involve non-Gaussian marginal distributions that are bounded, thus making the problems more challenging. These include distributions double-bounded, such as uniform, and singlebounded, such as exponential, Pareto and PBF, with the Pareto distribution also having a discontinuity at the origin (intermittence). The examples studied show how the problems of boundedness and discontinuity can be handled through simple post-processing procedures, thus achieving an overall satisfactory performance.

In conclusion, the method seems promising and expandable to several future research directions, such as multivariate stochastic modelling, downscaling, disaggregation, and stochastic modelling of two or more processes simultaneously, particularly in cases where time directionality is important (e.g. rainfall-runoff modelling at small time scales).

Stochastic simulation has recently acquired tremendous importance, as conventional energy sources are being replaced with renewables, whose nature is stochastic and thus their assessment needs stochastic tools. Its utility should now be appreciated more than ever, after various spectacular failures of aspirations to achieve satisfactory predictions of geophysical processes in deterministic terms, and after reconciliation with the fact that uncertainty is an intrinsic characteristic of nature, not subject to elimination.

Author Contributions: Conceptualization, DK and PD; methodology, DK.; software, DK validation, DK and PD.; formal analysis, DK and PD; investigation, DK and PD; resources, DK; data curation, DK; writing—original draft preparation, DK; writing—review and editing, DK and PD; visualization, DK; supervision, DK; literature review, DK and PD. Both authors have read and agreed to the published version of the manuscript

Funding: This research received no external funding but was conducted for scientific curiosity.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Discussions with Theano Iliopoulou, Federico Lombardo and Ioannis Tsoukalas have helped us in the method conceptualization.

Conflicts of Interest: The authors declare no conflict of interest. We expect eponymous reviewing to exclude any case that the reviewers could possibly have conflict of interest.

References

- Koutsoyiannis, D. Simple stochastic simulation of time irreversible and reversible processes. *Hydrological Sciences Journal* 2020, 65 (4), 536–551, doi: 10.1080/02626667.2019.1705302.
- Beven, K., Issues in Generating Stochastic Observables for Hydrological Models. *Hydrological Processes* 2021, doi.org/10.1002/hyp.14203p.e14203.
- 3. Box, G.E.; Jenkins, G.M. *Time Series Models for Forecasting and Control*. Holden Day, San Francisco, USA, 1970.
- 4. Stigler, S.M. Statistics on the Table: The History of Statistical Concepts and Methods. Harvard University Press, 2002.

- 5. Whittle, P. Hypothesis Testing in Times Series Analysis. PhD thesis, Almqvist & Wiksells, Uppsala, 1951.
- 6. Whittle, P. Tests of fit in time series. *Biometrika* **1952**, 39(3/4), 309-318.
- 7. Whittle, P. The analysis of multiple stationary time series. *Journal of the Royal Statistical Society B* **1953**, 15(1), 125-139.
- Dimitriadis, P.; Koutsoyiannis, D.; Iliopoulou, T.; Papanicolaou P. A global-scale investigation of stochastic similarities in marginal distribution and dependence structure of key hydrological-cycle processes. *Hydrology* 2021, 8 (2), 59, doi:10.3390/hydrology8020059.
- 9. Hosking, J.R.M. Fractional differencing, *Biometrika* 1981, 68(1), 165–176, doi: 10.1093/biomet/68.1.165.
- 10. Koutsoyiannis, D. A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. *Water Resources Research* **2000**, 36 (6), 1519–1533, doi: 10.1029/2000WR900044.
- 11. Koutsoyiannis, D. Generic and parsimonious stochastic modelling for hydrology and beyond. *Hydrological Sciences Journal* **2016**, 61 (2), 225–244, doi: 10.1080/02626667.2015.1016950.
- 12. Onof, C.; Chandler, R.E.; Kakou, A.; Northrop, P.; Wheater, H.S.; Isham, V. Rainfall modelling using Poisson-cluster processes: a review of developments. *Stoch. Environ. Res. Risk Assess.* **2000**, 14, 384–411, doi:10.1007/s004770000043.
- 13. Cowpertwait, P.; Isham, V.; Onof, C. Point process models of rainfall: Developments for fine-scale structure. *Proc. R. Soc. A* 2007, 463(2086), 2569–2588.
- 14. Kim, D.; Onof, C. A stochastic rainfall model that can reproduce important rainfall properties across the timescales from several minutes to a decade. *J. Hydrol.* **2020**, 589, 125-150, doi:10.1016/j.jhydrol.2020.125150.
- 15. Kossieris, P.; Makropoulos, C.; Onof, C.; Koutsoyiannis, D. A rainfall disaggregation scheme for sub-hourly time scales: Coupling a Bartlett-Lewis based model with adjusting procedures. *J. Hydrol.* **2018**, 556, 980–992, doi:10.1016/j.jhydrol.2016.07.015.
- Koutsoyiannis, D.; Onof, C.; Wheater, H.S. Multivariate rainfall disaggregation at a fine timescale. *Water Resources Research* 2003, 39 (7), 1173, doi: 10.1029/2002WR001600.
- 17. Northrop, P.J. A clustered spatial-temporal model of rainfall. Proc. R. Soc. London, Ser. A 1998, 454, 1875–1888.
- 18. 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.
- 19. Frechet, M., Sur les tableaux de correlation dont les marges son donnees. Ann Univ Lyon, Sect A 711 9, 53–77, 1951.
- 20. 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.
- 21. Nelsen, R.B., An Introduction to Copulas. Springer Series in Statistics, 2nd edition, 2006.
- 22. Lebrun, R.; Dutfoy, A. An innovating analysis of the Nataf transformation from the copula viewpoint. *Probabilistic Engineering Mechanics* **2009**, 24(3): 312-320.
- 23. Tsoukalas, I.; Makropoulos, C.; Koutsoyiannis, D. Simulation of stochastic processes exhibiting any-range dependence and arbitrary marginal distributions. *Water Resources Research* **2018**, 54 (11), 9484–9513, doi:10.1029/2017WR022462.
- 24. Eddington, A. The Nature of the Physical World. Cambridge University Press, Cambridge, UK, 1928.
- 25. Koutsoyiannis, D. Time's arrow in stochastic characterization and simulation of atmospheric and hydrological processes. *Hydrological Sciences Journal* **2019**, 64 (9), 1013–1037, doi: 10.1080/02626667.2019.1600700.
- 26. Koutsoyiannis, D. Simple stochastic simulation of time irreversible and reversible processes. *Hydrological Sciences Journal* **2020**, 65 (4), 536–551, doi: 10.1080/02626667.2019.1705302.
- 27. Vavoulogiannis, S.; Iliopoulou, T.; Dimitriadis, P.; Koutsoyiannis, D. Multiscale temporal irreversibility of streamflow and its stochastic modelling. *Hydrology* **2021**, 8 (2), 63, doi: 10.3390/hydrology8020063.
- 28. Lombardo, F.; Napolitano, F.; Russo, F.; Koutsoyiannis, D. On the exact distribution of correlated extremes in hydrology. *Water Resources Research* **2019**, 55 (12), 10405–10423, doi: 10.1029/2019WR025547.
- 29. Dimitriadis, P.; Koutsoyiannis, D. Stochastic synthesis approximating any process dependence and distribution. *Stochastic Environmental Research & Risk Assessment* **2018**, 32 (6), 1493–1515, doi: 10.1007/s00477-018-1540-2.
- 30. Koutsoyiannis, D. Knowable moments for high-order stochastic characterization and modelling of hydrological processes. *Hydrological Sciences Journal* **2019**, 64 (1), 19–33, doi: 10.1080/02626667.2018.1556794.
- 31. Koutsoyiannis, D. Stochastics of Hydroclimatic Extremes A Cool Look at Risk. ISBN: 978-618-85370-0-2, 333 pages, Kallipos, Athens, 2021.
- 32. Koutsoyiannis, D. Coupling stochastic models of different time scales. *Water Resources Research* 2001, 37 (2), 379–391, doi: 10.1029/2000WR900200.
- 33. Wold, H.O. A Study in the Analysis of Stationary Time-Series. PhD thesis, Almquist and Wicksell, Uppsala, 1938.
- 34. Wold, H.O. On prediction in stationary time series. *The Annals of Mathematical Statistics* **1948**, 19 (4), 558-567.
- 35. Thiele, T.N. Forelaesninger over Almindelig lagttagelseslaere: Sandsynlighedsregning og mindste Kvadraters Methode. C.A. Reitzel, Kjøbenhavn, 1889, <u>https://archive.org/details/forlaesingerove00thiegoog</u> (accessed on 18 May 2021).
- 36. Thiele, T.N. Om Iagttagelseslærens Halvinvarianter. Kgl. danske Videnskabernes Selskabs Forhandlinger 1899, 3, 135-141.
- 37. Hald, A. The Early History of the Cumulants and the Gram-Charlier Series. *International Statistical Review* **2000**, 68 (2), 137-153, doi: 10.1111/j.1751-5823.2000.tb00318.x.
- 38. Fisher, R. Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh, 1932.

- 39. Hotelling, H. Review of Statistical Methods for Research Workers, by R. A. Fisher. *Journal of the American Statistical Association* **1933**, 28 (183), 374-375, doi: 10.1080/01621459.1933.10502281.
- 40. Smith, P.J. A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *The American Statistician* **1995**, 49 (2), 217–218.
- 41. Koutsoyiannis, D. A random walk on water. *Hydrology and Earth System Sciences* **2010**, 14, 585–601, doi: 10.5194/hess-14-585-2010.
- 42. Kolmogorov, A.N. Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *Dokl. Akad. Nauk SSSR* **1940**, *26*, 115–118.
- 43. Kolmogorov, A.N. Wiener spirals and some other interesting curves in a Hilbert space. In *Selected Works of A. N. Kolmogorov Vol. 1, Mathematics and Mechanics*; Tikhomirov, V.M., Ed.; Kluwer: Dordrecht, The Netherlands, 1991; pp. 303-307.
- 44. Hurst, H.E. Long term storage capacities of reservoirs. Trans. Am. Soc. Civil Eng. 1951, 116, 776–808.
- 45. Koutsoyiannis, D. Entropy production in stochastics. Entropy 2017, 19 (11), 581, doi: 10.3390/e19110581.
- 46. Koutsoyiannis, D. Knowable moments for high-order stochastic characterization and modelling of hydrological processes. *Hydrological Sciences Journal* **2019**, 64 (1), 19–33, doi: 10.1080/02626667.2018.1556794.
- 47 Fernandez, B.; Salas, J.D. Periodic gamma autoregressive processes for operational hydrology. *Water Resour. Res.* **1986**, 22(10): 1385-1396.
- 48. Koutsoyiannis, D. Climate change, the Hurst phenomenon, and hydrological statistics. *Hydrological Sciences Journal* **2003**, 48 (1), 3–24, doi:10.1623/hysj.48.1.3.43481.
- 49. Koutsoyiannis, D.; Montanari, A. Statistical analysis of hydroclimatic time series: Uncertainty and insights. *Water Resources Research* **2007**, 43 (5), W05429, doi:10.1029/2006WR005592.
- 50. Dimitriadis, P.; Koutsoyiannis, D. Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst-Kolmogorov processes, *Stochastic Environmental Research & Risk Assessment* **2015**, 29 (6), 1649–1669, doi: 10.1007/s00477-015-1023-7.
- 51. Koutsoyiannis, D. An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence. *Water Resources Research* **2006**, 42 (1), W01401, doi: 10.1029/2005WR004175.
- 52. McCullagh P; Kolassa, J. Cumulants, *Scholarpedia* 2009, 4 (3), 4699, <u>http://www.scholarpedia.org/article/Cumulants</u> (accessed on 18 May 2021).