Stochastic similarities between the microscale of turbulence and hydrometeorological processes

Panayiotis Dimitriadis*, Demetris Koutsoyiannis and Panos Papanicolaou

*Corresponding author; email: pandim@itia.ntua.gr

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

Abstract

Turbulence is considered to generate and drive most geophysical processes. The simplest case is the isotropic turbulence. In this paper, the most common three-dimensional power-spectrum-based models of isotropic turbulence are studied in terms of their stochastic properties. Such models often have a high-order of complexity, lack in stochastic interpretation and violate basic stochastic asymptotic properties, such as the theoretical limits of the Hurst coefficient, in case that Hurst-Kolmogorov behaviour is observed. A simpler and robust model (which incorporates self-similarity structures, e.g. fractal dimension and Hurst coefficient) is proposed using a climacogram-based stochastic framework and tested over high resolution observational data of laboratory scale as well as hydrometeorological observations of wind speed and precipitation intensities. Expressions of other stochastic tools like the autocovariance and power spectrum are also produced from the model and show agreement with data. Finally, uncertainty, discretization and bias related errors are estimated for each stochastic tool, showing lower errors for the climacogram-based ones and larger for power-spectrum ones.

Keywords: isotropic-stationary turbulence; hydrometeorological processes; stochastic modelling; climacogram; power spectrum; uncertainty-bias

1. Introduction

Turbulence originates from the Greek word 'τύρβη' (cf. '…τὴν τύρβην ἐν ᾗ ζῶμεν':'…for the turbulence in which we live', Isokrates, 15.130) which means disorder, confusion, turmoil. Turbulence is considered to generate and drive most geophysical processes, e.g. wind turbulence giving birth and spatiotemporal variability in cloud rainfall (cf. Falkovich et al. 2002), yet it is regarded as mystery within classical physics (McDonough 2007 ch. 1). Studying turbulent phenomena is of high importance for hydrology (e.g. Mandelbrot and Wallis 1968, Rinaldo 2006) as the microscopic processes (related to turbulence) can help understand the macroscopic ones (related to hydrology), since they enable the recording of very long time-series and with a high resolution, a rare case for hydrological processes (cf. Koutsoyiannis 2014). The simplest case of turbulent state (in terms of mathematical calculations) is the stationary, isotropic and homogeneous turbulence. While this is a physical phenomenon that has been recognized hundreds of years ago, still there is no universally agreed mathematical definition for the so-called 'turbulent state' (Tessarotto and Asci 2010). Leonardo da Vinci tried to give a definition 500 years ago, based on his observations that water falling into a sink forms large eddies as well as rotational motion (cf. Richter 1939). Interestingly, Heisenberg (1948) commented on the definition of turbulent state of flow that it is just the result of infinite degrees of freedom developed in a liquid flowing without friction and thus, by contrast, laminar flow is a state of flow with reduced degrees of freedom caused by the viscous action. In 1880, Reynolds introduced one of the most important dimensionless parameters in fluid mechanics, the ratio of momentum over viscous forces which is called Reynolds number ever since. Based on this dimensionless parameter, it was observed that irrotationality in the streamlines occurred for values much greater than 1 and led to somehow confine the occurrence of turbulence to Reynolds number values greater than approximately 1000 to 2000. Richardson (1922) introduced the idea of turbulence 'energy cascade' by stating that turbulent motion, powered by the kinetic energy, is first produced at the largest scales (through eddies of size comparable to the characteristic length scale of the natural process) and then to smaller and smaller ones, until is dissipated by the viscous strain action. Taylor (1935) was the first to use stochastic tools to study this phenomenon modelling turbulence by means of random variables rather than deterministic ones. Following this idea, Kolmogorov (1941a-c) managed to derive the famous '5/3' law (K41 theory) using the Navier-Stokes equations. That law describes the energy dissipation rate from larger to smaller turbulence scales within the inertial wavenumber sub-range, with the power spectrum no longer dependent on the eddy size and fluid viscosity. Since then, many scientists (including Von Karman 1948, Heisenberg 1948, Kraichnan 1959, Batchelor 1959, Pope 2000), have significantly contributed to the current power-spectrum-based models of turbulence.

A general view of the stochastic approach of stationary and isotropic turbulence (in which the random variables describing turbulence have the same statistical properties in all directions) can be seen in many text books, e.g. Pope (2000). In this paper, we focus on the investigation of the second-order statistics (e.g. power spectrum) and the preservation of the marginal probability density function (pdf). We are mainly interested in the local and global stochastic properties of a process, by calculating its fractal dimension and by examining whether it exhibits HK behaviour, respectively. Furthemore, we investigate the stochastic properties of the most common three-dimensional power-spectrum-based models of stationary and isotropic turbulence in time domain and we detect some model weaknesses despite their widespread use. A simpler and more robust model, which incorporates both fractal and Hurst-Kolmogorov (HK) possible behaviours, is proposed using a second-order stochastic framework based on the concept of climacogram. This model is tested over high resolution nearly isotropic observational data of laboratory scale. Moreover, we show that the same model can be used for small-scale hydrometeorological processes generated by turbulence such as atmospheric wind speed and precipitation intensities. Expressions of other stochastic tools such as the autocovariance and power spectrum are also produced directly from the model and are in agreement with data. Finally, uncertainty, discretization and bias related errors are estimated for each stochastic tool, showing, in general, lower errors for the climacogram-based model and larger ones for power-spectrum based ones. It is noted that the HK process corresponds to Fractional Gaussian Noise (cf. Mandelbrot and Wallis 1968) and is named after Hurst (1951), who first detected the long-term behaviour in geophysical time-series and Kolmogorov (1940) who first introduced the mathematical form of the process (cf. Koutsoyiannis 2011a).

2. Definitions and notations

Stochastic modelling and probabilistic approaches have been proven useful in the investigation of processes that resist a deterministic description, such as turbulence (e.g. Kraichnan 1991 ch. 1, Frisch, 2006 ch. 3, McDonoug, 2007 ch. 1, Koutsoyiannis 2014). Using stochastic mathematical processes one can represent, and thus interpret, a natural process based on its statistical properties whose values can be estimated through stochastic tools such as autocovariance-based ones defined in the equations below:

$$c(\tau) := \mathrm{Cov}\left[\underline{x}(t), \underline{x}(t+\tau)\right] \tag{1}$$

$$v(\tau) := c(0) - c(\tau) \tag{2}$$

$$s(w) := 4\int_0^\infty c(\tau)\cos(2\pi w\tau)\,\mathrm{d}\tau \tag{3}$$

where $\underline{x}(t)$ is the continuous time process (underscore denotes a random variable), $c(\tau)$ is the autocovariance function, $v(\tau)$ the variogram (else known as 2nd structural function), $s(w)$ the power spectrum and τ, w the continuous time lag and frequency, respectively (see in Appendix for details).

Other stochastic tools can be based on the climacogram (e.g. Koutsoyiannis 2013a), which is defined as the (plot of) variance of the averaged process $\frac{1}{m}\int_0^m \underline{x}(t)\mathrm{d}t$ (assumed stationary) *vs* averaging time scale m and is denoted as $\gamma(m)$:

$$\gamma(m) := \frac{\mathrm{var}\left[\int_0^m \underline{x}(t)\mathrm{d}t\right]}{m^2} \tag{4}$$

The climacogram is useful to measure the variance of a process among scales (the kinetic energy, in case the variable under consideration is the velocity), and has many advantages in stochastic model building, namely small statistical as well as uncertainty errors (Dimitriadis and Koutsoyiannis 2015). It is also directly linked to the autocovariance function by the following equations (Koutsoyiannis 2013a):

$$\gamma(m) = 2\int_0^1 (1-x)c(xm)\mathrm{d}x \tag{5}$$

$$c(\tau) = \frac{\partial^2(\tau^2\gamma(\tau))}{2\partial\tau^2} \tag{6}$$

A climacogram-based spectrum (CBS), else known as the 'pseudospectrum', for comparison with the classical power spectrum, can be also defined as (Koutsoyiannis 2013a):

$$\psi(m) := \frac{2\gamma(1/w)}{w}\left(1 - \frac{\gamma(1/w)}{\gamma(0)}\right) \tag{7}$$

Furthermore, we introduce here, a climacogram-based variogram (CBV) for comparison with the classical variogram:

$$\xi(m) := \gamma(0) - \gamma(m) \tag{8}$$

Note that both CBS and CBV include the process variance at scale 0, i.e. $\gamma(0)$ and thus, they are applied only after a stochastic model is set.

All the above stochastic tools definitions and expressions in discrete time as well as widely used estimators, estimations (based on the latter estimators) and expected values, can be found in Appendix.

3. Most common stochastic models of stationary and isotropic turbulence

It is noted that the log-log derivative (LLD) is an essential concept in turbulence as it can identify possible scaling behaviour related to asymptotic coefficients (e.g. fractal dimension and Hurst coefficient). The LLD of any function $f(x)$ is defined as:

$$f^{\#}(x) := \frac{\mathrm{d}\ln(f(x))}{\mathrm{d}\ln x} = \frac{x}{f(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x} \tag{9}$$

and for the finite logarithmic derivative of $f(x)$, e.g. in case of discrete time process, we choose the backward log-log derivative, i.e.:

$$f^{\#}(x_i) \coloneqq \frac{\ln\left(f(x_i)/f(x_{i-1})\right)}{\ln(x_i/x_{i-1})} \tag{10}$$

Based on Gneiting et al. (2012) analysis, the fractal dimension (F) can be defined as (cf. Beran et al. 2013 ch. 3.6):

$$F \coloneqq N + 1 - \frac{1}{2}\lim_{\tau \to 0} \xi^{\#}(\tau) \tag{11}$$

where N the dimension of the field (e.g. N=1 for 1D velocity field).

Based on Beran et al. (2013 ch. 1.3) analysis, the Hurst coefficient (H) can be defined as:

$$H \coloneqq 1 + \frac{1}{2}\lim_{m \to \infty} \gamma^{\#}(m) \tag{12}$$

3.1 Commonly used processes

Following the stochastic framework in Section 2 (and in Appendix), we derive in Table 1, the 1D and 3D isotropic power spectra as well as their LLD's, for a Markovian process, a special case of a powered-exponential process (e.g. Yaglom 1987 ch. 10, Gneiting et al. 2012) and a generalized HK (gHK) process (cf. Dimitriadis and Koutsoyiannis 2015), which the latter behaves as Markovian-like for small scales and HK-like for large ones. These positively-correlated mathematical processes enclose possible asymptotic behaviours in large and small scales. In particular, a positively-correlated natural process may approach zero or infinite scale, by a powered-exponential (e.g. Markovian process) or a power-type (e.g. HK process) rise or decay, respectively. The 1D power spectrum and the 3D one, denoted as $s_{3\mathrm{D}}(\boldsymbol{w})$, are related by (Batchelor 1959 p. 50, Pope 2000 pp. 226-227, Kang et al. 2003):

$$s(w) = \int_1^{\infty} \frac{x^2-1}{x^3} s_{3\mathrm{D}}(\|\boldsymbol{w}\| x)\, \mathrm{d}x \tag{13}$$

$$s_{3\mathrm{D}}(w) = \frac{w^3}{2} \frac{\partial\left(\frac{1}{w}\frac{\partial(s(w))}{\partial w}\right)}{\partial w} \tag{14}$$

where $\boldsymbol{w}$ is the isotropic 3D frequency vector, with $\|\boldsymbol{w}\| = w \geq 0$.

As mentioned above, the most common used model for stationary and isotropic turbulence consists of the work of many scientists. Combining them into one equation, the power spectrum of isotropic and stationary turbulence can be written as (Pope 2000 pp. 232-233, Cerutti and Meneveau 2000, Kang et al. 2003):

$$s_{3\mathrm{D}}(w) = f_{\mathrm{E}}(w, c_{\mathrm{E}}, p) f_{\mathrm{I}}(w, c_{\mathrm{I}}) f_{\mathrm{D}}(w, c_{\mathrm{D}}) \tag{15}$$

where c_{E}, c_{I}, c_{D} and p are model parameters (see Pope 2000, pp. 233 for description) and from the work of Von Karman (1948), for the from the work of Von Karman (1948), for the energy containing eddies (large scales):

$$f_{\mathrm{E}}(w, c_{\mathrm{E}}, p) = \left(\frac{w}{\sqrt{w^2 + c_{\mathrm{E}}}}\right)^{\frac{5}{3}+p} \tag{16}$$

combined with the work of Kolmogorov (1941a-c) for the inertial range (intermediate scales):

$$f_{\mathrm{I}}(w, c_{\mathrm{I}}) = c_{\mathrm{I}} w^{-\frac{5}{3}} \tag{17}$$

and from the work of Kraichnan (1959) for the dissipation range (small scales):

$$f_{\mathrm{D}}(w, c_{\mathrm{D}}) = \mathrm{e}^{-wc_{\mathrm{D}}} \tag{18}$$

Table 1: 1D and 3D power spectrum for Markovian, powered-exponential and gHK processes as well as their LLD's (estimated from equation 9), where λ is the parameter related to the true variance of the process, q the scale parameter and b is related to the power-type behaviour of the process.

Markovian		Powered-exponential special case		gHK	
$c(\tau) = \lambda \mathrm{e}^{-\lvert\tau\rvert/q}$	(19)	$c(\tau) = \lambda \mathrm{e}^{-(\tau/q)^2}$	(20)	$c(\tau) = \lambda \dfrac{(1-b)(2-b)}{(1+\lvert\tau\rvert/q)^b}$ with $b \in (0,2)$	(21)
$s(w) = \dfrac{4\lambda q}{1+4\pi^2 q^2 w^2}$ with $\lim_{w\to 0} s^\# = 0$ and $\lim_{w\to\infty} s^\# = -2$	(22)	$s(w) = \dfrac{\lambda q\sqrt{\pi}}{2}\mathrm{e}^{-(qw\pi)^2}$ with $s^\#(w) = -2(qw\pi)^2$, $\lim_{w\to 0} s^\# = 0$ and $\lim_{w\to\infty} s^\# = -\infty$	(23)	$\lim_{w\to 0} s \sim w^{b-1}$ with $\lim_{w\to 0} s^\# = b-1$	(24)
				$\lim_{w\to\infty} s \sim w^{-2}$ with $\lim_{w\to\infty} s^\# = -2$	(25)
$s_{3\mathrm{D}}(w) = \dfrac{4\lambda q(2\pi q w)^4}{(1+4\pi^2 q^2 w^2)^3}$ with $\lim_{w\to 0} s_{3\mathrm{D}}{}^\# = 4$ and $\lim_{w\to\infty} s_{3\mathrm{D}}{}^\# = -2$	(26)	$s_{3\mathrm{D}}(w) \sim q^5 w^4 \mathrm{e}^{-(qw\pi)^2}$ with $s^\#(w) = 4 - 2(qw\pi)^2$ $\lim_{w\to 0} s_{3\mathrm{D}}{}^\# = 4$ and $\lim_{w\to\infty} s_{3\mathrm{D}}{}^\# = -\infty$	(27)	$\lim_{w\to 0} s_{3\mathrm{D}} \sim w^{b-1}$ with $\lim_{w\to \mathbf{0}} s_{3\mathrm{D}}{}^\# = b-1$	(28)
				$\lim_{w\to\infty} s_{3\mathrm{D}} \sim w^{-2}$ with $\lim_{w\to\infty} s_{3\mathrm{D}}{}^\# = -2$	(29)

3.2 Stochastic properties of large-scale range

For the 3D and 1D (derived from the 3D one) power spectra at the energy containing range, we have that:

$$\lim_{w\to \mathbf{0}} s_{3\mathrm{D}} = \lim_{w\to 0} s \sim w^p \tag{30}$$

where Von Karman (1948) suggests $p = 4$ (or else known as 'Batchelor turbulence', cf. Davidson 2000), while other works result in different values, e.g. Saffman (1967) suggests $p = 2$.

There are many arguments about the proper value of the p parameter and its relation to the Loitsyansky integral which controls the rate of decay of kinetic energy (cf. Davidson 2000). The main debate is whether points at a large distance in stationary, isotropic and homogeneous turbulent flow are statistically independent or show a correlation that decays either exponentially (e.g. Von Karman model for wind gust, cf. Wright and Cooper 2007 ch. 16.7.1; Faisst and Eckhardt 2004, Avila et al. 2010 and Kuik et al. 2010, models for pipe flow) or with a power-type law (see below for several examples).

Towards the stochastic properties of the aforementioned equation, we can see from Table 1 that the case $p = 2$ does not correspond neither to exponential (Markovian or powered-exponential) nor to power-type (i.e. HK) decay of autocovariance. Hence, this model cannot be applied to asymptotic zero frequencies (or infinite scales). Interestingly, the case $p = 4$ can be interpreted by a Markovian (equation 26) or a special case of the powered-exponential (equation 27) decay of autocovariance. However, this case also excludes the HK behaviour, i.e. autocovariance long-range dependence (e.g. equation 21), where p now equals $b - 1$ and is bounded to [-1, 1].

Although the aforementioned models do not include a possible power-law decay of autocovariance (i.e. HK behaviour), several works show strong indication that turbulence natural processes can exhibit HK behaviour rather than Markovian. Such works are reported by e.g., Nordin et al. (1972) for laboratory turbulent flume and turbulent river velocities, Helland and Van Atta (1978) for grid turbulence velocities, Goldstein and Roberts (1995) for magneto-hydrodynamic turbulent solar wind, Chamorro and Potre-Agel (2009) for wind turbulent wakes and grid-turbulence, Dimitriadis and Papanicolaou (2012) and Charakopoulos et al. (2014a,b) for turbulent buoyant jets, Koutsoyiannis (2013b) for grid turbulence. Koutsoyiannis (2011b) has also shown that entropy maximization results in HK dynamics at asymptotic times (zero or infinity) under the constraints of mean, variance and autocovariance of lag one preservation.

We believe that the reason a possible HK behaviour is not detected in geophysical processes (which are often characterized by lack of measurements), is that mathematical smoothing techniques are applied (e.g. windowing or else Welch approaches, regression analysis, wavelet techniques, see other examples in Stoica and Moses 2004 ch. 2.6). Particularly, application of windowing techniques to any stochastic tool can be misleading since they eliminate a portion (depending on the type and length of the window applied) of the time-series' variance (which often is incorrectly attributed to 'noise', cf. Koutsoyiannis, 2010). This elimination can lead to process' misrepresentation in case of significant effects of discretization, small and/or finite record length and bias (examples of applications to the power spectrum can be seen in e.g. Lombardo et al. 2013). An example of smoothing out the HK behaviour by applying the Welch approach with a Bartlett window and no segment-overlapping to an observed time-series, is shown in Fig. 1(a). Even though the smoothing technique decreases the power spectrum variance, it also causes low frequency loss of information (e.g. see other examples in Dimitriadis et al., 2012). This loss of information may cause a process misinterpretation, as illustrated in Fig. 1(b), where the 1D autocorrelation function (derived from the 3D power spectrum model in equation 15) exhibits a Markovian-like decay, while the empirical one (derived from the windowed empirical power spectrum partitioned into 10^3 segments) exhibits an HK behaviour. Also, this smoothing technique should be used in caution in strong-correlated processes, as increasing the number of partitioned segments will also cause an increase in their cross-correlation (Fig. 1a). Finally, processes with HK behaviour have usually large bias and in case this is not included in the model, the empirical autocovariance's rapid decay in large scales (or equivalently lags) may be erroneously interpreted as short-range dependence (Fig. 1b).

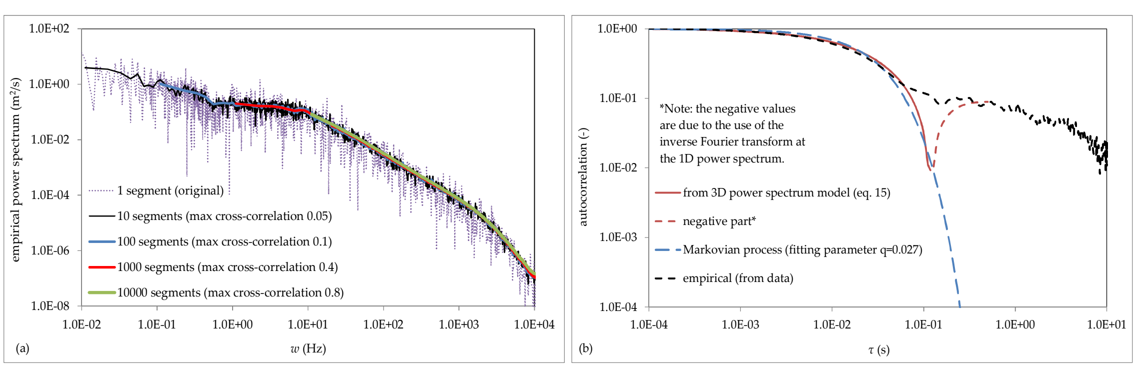

Fig. 1: (a) Example of loss of low frequency information caused by the application of the windowing technique, in a time-series provided by the Johns Hopkins University (see also in Section 4 for more details on the dataset) as well as the maximum cross correlations between the partitioned segments;

(b) 1D autocorrelation function derived from the 3D power spectrum model in equation 15 (with parameters based on the fitting of the windowed 1D power spectrum with 1000 segments in Fig. 1(a): $c_\mathrm{E} = 2.5\ \mathrm{m}^{-2}$, $p = 4$, $c_\mathrm{I} = 13.0\ \mathrm{m}^3/\mathrm{s}^2$, $c_\mathrm{D} = 2 \times 10^{-4}$ m); a Markovian autocorrelation function, i.e. $e^{-(\tau/q)}$, for reasons of comparison; and the corresponding (to the windowed 1D power spectrum with 1000 segments in Fig. 1a) empirical autocorrelation function.

To incorporate possible HK behaviour in the model, we may assume an autocovariance power-type decay at large scales, where the 3D and 1D power spectra at asymptotically zero frequency are of the form w^{b-1} (Table 1), with b bounded to (0,2), for positively correlated processes ($0.5 < H < 1$), negatively-correlated processes ($0 < H < 0.5$) and for a process with a random decay in large scales ($H = 0.5$), with H the Hurst coefficient ($H = 1 - b/2$, from equation 12).

3.3 Stochastic properties of small-scale range

Similarly, for the 3D and 1D power spectra at the dissipation range, we have that:

$$\lim_{w\to\infty} s_{3\mathrm{D}}(w) = \lim_{w\to\infty} s(w) \sim \mathrm{e}^{-w} \qquad (31)$$

This results in autocovariance function of the form:

$$c(\tau) \sim \frac{1}{\tau^2+1} \qquad (32)$$

which corresponds to Wackernagel (1995) process (he also refers to it as autocovariance-based Cauchy-class process resembling the Cauchy probability function). A generalized expression of this process can be found in Gneiting (2000), which we will refer to it as the Gneiting process (its analytical expressions are shown in Section 4.2). For small lags this process behaves like (e.g. Gneiting and Schlather 2004):

$$\lim_{\tau\to 0} c(\tau) \sim 1 - \tau^2 \sim e^{-\tau^2} \qquad (33)$$

which corresponds to the special case of a powered-exponential process in Table 1. Note, that this process corresponds to $H = 0$ (based on the definition in equation 12), if applied to large scales.

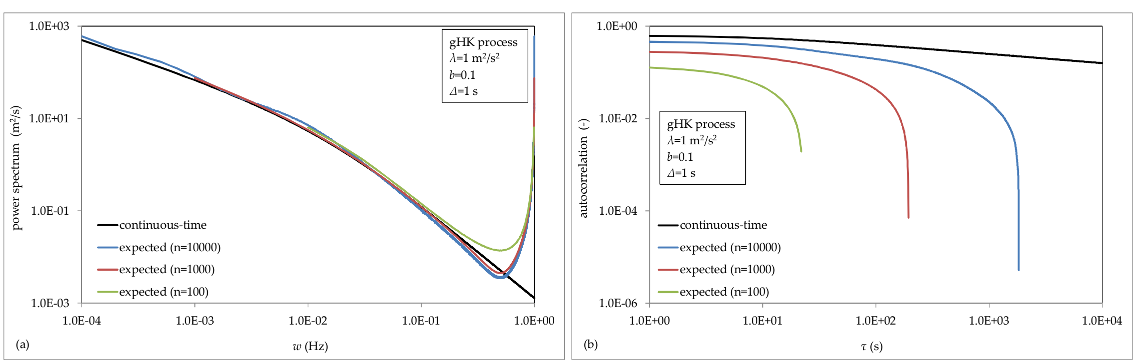

Fig. 2: (a) Power spectra and (b) corresponding autocovariances, in continuous time as well as their expected values, with varying number of records (denoted as n) of a gHK process. The expected autocovariance and power spectrum are estimated from equation (A17) and (A25), respectively (see Appendix).

Other models for the dissipation range are of the form of a powered-exponential power spectrum process (e.g. Cerutti and Meneveau 2000) which may result from a powered-exponential autocovariance function (Table 1). However, there is evidence that these models cannot interpret the frequently observed spike in the high frequency power spectrum (e.g. Cerutti and Meneveau 2000, Kang et al. 2003). This is usually ignored and attributed to instrumental noise. Here, we show that this spike may appear in HK processes and is due to discretization and bias errors, in case the shape parameter q/Δ takes large values (Fig. 3).

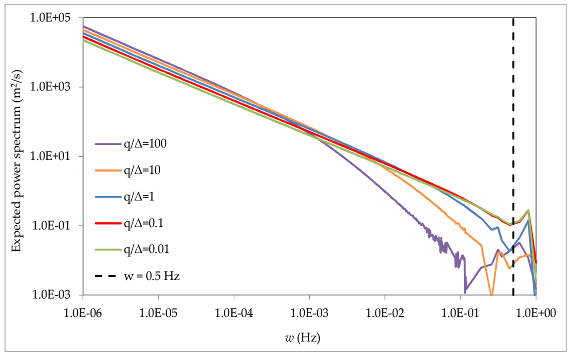

Fig. 3: Expected power spectra (estimated from equation A25) of a gHK process, with varying q/Δ (where Δ the sampling time interval, see in Appendix for its relation to the expected value of a stochastic tool).

3.4 Stochastic properties of intermediate-scale range

From Table 1, one may observe that the power spectrum asymptotic LLD's from different processes, are often coincident with each other. For example, for both a Markovian and a gHK process with b =1, the power spectrum LLD is 0 for the low frequency tail and -2 for the high frequency one. This may be confusing and result in misinterpretation of the natural process. A solution to this may be to incorporate additional stochastic tools in the analysis as shown in Section 4. For the aforementioned example, if the autocovariance function asymptotic properties (local and global ones) are analyzed, one can decide upon a powered-exponential lag decay (e.g. a Markovian process) and a power-type one (e.g. a gHK process). At the same basis, when a power-type behaviour appears in the intermediate frequencies of a power spectrum (e.g. in case of a -5/3 LLD), it may be misleading to interpret it as a power-law function (and thus, a power-type autocovariance decay, as shown in Table 1), because this can result from different kind of processes which they do not have power-type expressions for the intermediate scale-range. An illustrative example is shown in Fig. 4, where the -5/3 LLD in the intermediate frequencies of the power spectrum results from a simple combination of a Markovian and a gHK process, both of which have a purely stochastic interpretation and they do not include power-type laws in the intermediate frequency-range.

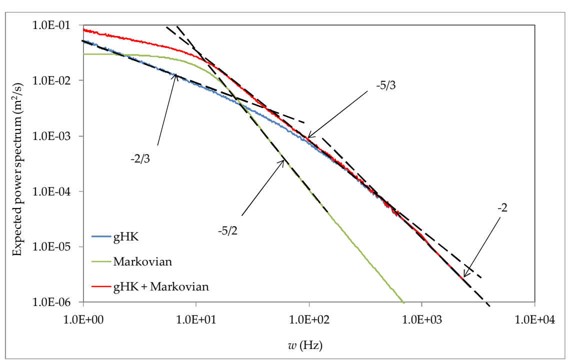

Fig. 4: Expected power spectrum (estimated from equation A25) resulted from a combination of a Markovian and a gHK process (with parameters same as in the application of section 4.1 and N=10^4).

Note also, that the Kolmogorov (1941a-c) power-type power spectrum refers only to intermediate frequencies and should not be applied arbitrarily for low frequencies too, as the corresponding autocovariance asymptotic large-scale behaviour, i.e. $c(\tau) \sim \tau^{\frac{5}{3}-1}$, gives an invalid (based on equation 12) $H = 4/3 > 1$.

4. Proposed model and applications

In the previous section, we present several limitations concerning the stochastic properties of proposed turbulent models from literature. Specifically, we see that they only include exponential decay in the energy containing area and thus, completely excluding possible HK behaviour. They also, describe the dissipation area decay with only a specific case of a powered-exponential process and thus, leaving out all other possible types of decay. Moreover, they interpret a possible power-type-like intermediate area (of the power spectrum) with power-type behaviour (and particularly, only that of the K41 theory) which can also result from intermediate non power-type processes (as shown in Fig. 4). Furthermore, these models are based only on the power spectrum stochastic tool (causing possible misinterpretation in other tools, e.g. climacogram, autocovariance) and on multiple processes multiplication (which may cause numerical difficulties in stochastic generation). Since turbulence generates and drives most of geophysical processes, we expect geophysical processes to exhibit similar types of decay in small and large scales. Hence, a more robust, flexible and parsimonious model is required that can incorporate all the aforementioned microscale and macroscale behaviours linking turbulence to hydrology. Here, we choose the ergodic stochastic model in Table 2, which consists of two independent processes, that of a powered exponential (controlling the small scales and fractal behaviour, cf. Gneiting et al. 2012) and a gHK (controlling the large scales and HK behaviour, cf. Dimitriadis and Koutsoyiannis, 2015), which are combined in a way to exhibit the desired expected LLD in the intermediate scales. This model can describe all linear combinations of powered-exponential and HK processes, including the often observed intermediate quick drop of all the stochastic tools (see Section 4.1, 4.2 and 4.3, for an example in grid turbulence, wind and precipitation process). This particular drop may be due to the interference of boundaries and/or the existence of multiple periodic functions, as for example in case of combinations of HK with cyclostationary processes (cf. Markonis and Koutsoyiannis 2013). Furthermore, although the proposed model results in a complicated power spectrum expression (equation 37), it provides simpler expressions for the other tools if compared to the most common model described in Section 3 (which has no analytical expressions for all tools except for the power spectrum). Finally, the proposed model is also justified by the maximization of entropy production in logarithmic time (abbreviated EPLT), a term introduced and defined by Koutsoyiannis (2011c) as the LLD of entropy. Particularly, Koutsoyiannis (2015) showed that the powered-exponential process has the largest EPLT for the microscale range (time-scale tending to zero) and the HK process has the largest EPLT for the macroscale range (time-scale tending to infinity). Hence, the maximization of EPLT can result from a combination of both processes.

Table 2: Autocovariance, variogram, climacogram, CBV, CBS and power spectrum mathematical expressions of the stochastic model, consisted of two independent processes in continuous time, that of a powered exponential and a gHK.

Type	Stochastic model	
Autocovariance*	$c(\tau) = \lambda_1 \mathrm{e}^{-(\lvert\tau\rvert/q_1)^{a}} + \lambda_2 (1 + \lvert\tau\rvert/q_2)^{-b}$	(34)

Climacogram
$$\gamma(m) = \frac{2\lambda_1\left(\frac{m}{q_1}\Gamma_l(1/a,(m/q_1)^{\alpha}) - \Gamma_l(2/a,(m/q_1)^{\alpha})\right)}{(m/q_1)^2} + \frac{2\lambda_2((m/q_2+1)^{2-b} - (2-b)m/q_2 - 1)}{(1-b)(2-b)(m/q_2)^2} \tag{35}$$

Variogram
$$v(\tau) = \lambda_1 + \lambda_2 - c(\tau) \tag{36}$$

Power spectrum**
$$s(w) = \mathrm{ICF}\left[\lambda_1 \mathrm{e}^{-(|\tau|/q_1)^{\alpha}}\right] + \frac{4\lambda_2 {q_2}^b\,\Gamma(1-b)\mathrm{Sin}\left(\frac{\pi b}{2} + 2q_2\pi|w|\right)}{(2\pi|w|)^{1-b}} - \frac{4\lambda_2 q_2\ {}_1\mathrm{F}_2\left[1;1-\frac{b}{2},\frac{3}{2}-\frac{b}{2};-\pi^2{q_2}^2w^2\right]}{1-b} \tag{37}$$

CBV
$$\xi(m) = \lambda_1 + \lambda_2 - \gamma(m) \tag{38}$$

CBS
$$\psi(w) = \frac{2\gamma(w)}{w}\left(1 - \frac{\gamma(w)}{\lambda_1 + \lambda_2}\right) \tag{39}$$

* $\lambda_2 = \lambda(1-b)(2-b)$, with λ a parameter related strictly to the process' variance.

** Since the inverse cosine Fourier (ICF) transform of the powered-exponential function and the hyper-geometric function ${}_1\mathrm{F}_2$ have not an analytical form, this cannot be written in a closed expression and numerical algorithms must be used.

4.1 Application to small-scale grid turbulence

In this section, we show the stochastic analysis of a grid-turbulence process based on a large open access dataset (http://www.me.jhu.edu/meneveau/datasets/datamap.html), provided by the Johns Hopkins University. Microscale turbulence description has many applications in hydrometeorological processes which often lack small scale measurements (cf. Koutsoyiannis 2011c), thus introducing limitations in the fitted models (e.g. the fractal dimension of the process cannot be estimated based on the definition of equation 11). An illustrative example of an application to atmospheric wind speed is shown in Section 4.2.

Here, we only consider the longitudinal wind velocity dataset along the flow direction since the other two components are limited by the experiment's construction boundaries. This dataset consists of 40 time-series (Fig. 5a), measured by X-wire probes placed downstream of the grid (Kang et al. 2003). The first 16 time-series correspond to velocities measured at transverse points abstaining r = 20M from the source, where M = 0.152 m is the size of the grid. The next 4 time-series correspond to distance r = 30M, the next 4 to 40M and the last 16M to 48M. For details regarding the experimental setup and datasets see Kang et al. (2003). All time-series are considered to be stationary with a nearly-Gaussian probability density function (see in Fig. 5c), are nearly isotropic with isotropy ratio 1.5 (Kang et al. 2003) and very long (each contains n = 36×10^6 data points), covering all three aforementioned scale ranges of equation (15). Moreover, the sampling time interval, denoted as D, is considered small (2.5 μs), therefore equality $D = \Delta$, where Δ (≤ D) the instrument response time, can be assumed valid. In Appendix, we noted that if D is small the differences between stochastic processes in discretized time with $\Delta > 0$ and $\Delta \approx 0$ are also expected to be small. Finally, following the same analysis of Dimitriadis and Koutsoyiannis (2015), the expected value of each examined stochastic tool can be roughly estimated as the average value of all 40 time-series (Fig. 6a-g), after homogenization is applied (the marginal variance of the process is estimated approximately 2.272 m^2/s^2). Additionally, we choose the 38$^{\text{th}}$ time-series for the empirical one, after observing that is the closest one to each stochastic tool's averaged value (Fig. 6h). Since we expect this to be near to the process expected values, it can help us test the validity of the stochastic model. Modelling phenomena such as intermittency (which is related to high-order derivatives, c.f. Kang et al. 2003, Batchelor and Townsend 1949) as well as preservation of high order moments (which are often characterized by high uncertainty, cf. Lombardo et al. 2014) deviate from the purpose of this paper. In this paper, we are mainly interested in the local and global 2nd order stochastic properties of the process, by calculating the process fractal dimension and by examining whether the process exhibits HK behaviour, respectively.

As we have already mentioned, the velocity field is not homogeneous and the root-mean-square (rms) velocity components (i.e. standard deviations of velocity) are decreasing with the distance from the grid (Fig. 5b). To make data homogeneous, we normalize each time-series by subtracting the mean $\mu_{\mathrm{t}}(r)$ and dividing by the standard deviation $\sigma_{\mathrm{t}}(r)$, both estimated from the equations of the fitted curves in Fig. 5(b):

$$\sigma_{\mathrm{t}}(r) = 4.16(r + 0.3)^{-0.657} \tag{40}$$

$$\sigma_{\mathrm{t}}(r)/\mu_{\mathrm{t}}(r) = 0.859r + 3.738 \tag{41}$$

where r is the distance from the grid. Note that coefficient 0.3 in equation (40) has been added for consistency reasons, so that the variance is finite at distances near the grid.

We also observe that the pdf of the time-series are not exactly Gaussian, since for example the empirical skewness is approximately equal to 0.2 (Fig. 5c and 5d). Here, we propose a normalization scheme by separating the empirical pdf to multiple segments and then approximating them with multiple Gaussian distributions:

$$f_{\mathrm{t}}(u) = \begin{cases} \mathrm{N}(\mu_1, \sigma_1), & -\infty < u_1 \leq h_1 \\ \mathrm{N}(\mu_2, \sigma_2), & h_1 < u_2 \leq h_2 \\ & \dots \\ \mathrm{N}(\mu_o, \sigma_o), & h_{o-1} < u_o < \infty \end{cases} \tag{42}$$

where $f_{\mathrm{t}}(u)$ is the model pdf of the velocity u, $\mathrm{N}(\mu_l, \sigma_l)$ is a Gaussian pdf for the u_l branch of the empirical pdf (consisted of all quantiles $h_{l-1} < u_l \leq h_l$), with l varying from 1 to o (with $h_0 \to -\infty$ and $h_o \to \infty$) and with o representing the number of branches we separate the empirical pdf.

The μ_l and σ_l parameters can be calculated by simply fitting $\mathrm{N}(\mu_l, \sigma_l)$ to the empirical pdf of the quantiles within the l segment (subject to the constraints that the cdf and pdf values between the multiple Gaussian functions are equal). Specifically, if the l segment consists of only two quantiles, u_1 and u_2, and with $F_1$and F_2, the empirical cumulative distribution function (cdf) at these points, then the above parameters are obviously equal to:

$$\mu_l = u_1 - \sigma_l \sqrt{2}\,\mathrm{erf}^{-1}(2F_1 - 1) \tag{43}$$

$$\sigma_l = \frac{u_2 - u_1}{\sqrt{2}\left(\mathrm{erf}^{-1}(2F_2 - 1) - \mathrm{erf}^{-1}(2F_1 - 1)\right)} \tag{44}$$

with erf^{-1} the inverse of the error function.

Then, we can easily transform $u \sim f_{\mathrm{t}}$ to $u_{\mathrm{n}} \sim \mathrm{N}(0,1)$, by simply subtracting from each set of quantiles $(h_{l-1} < u_l \leq h_l)$ the mean μ_l and then dividing with the standard deviation σ_l. Furthermore, the reverse transformation scheme from a variable $u_{\mathrm{n}} \sim \mathrm{N}(0,1)$ to $u_{\mathrm{r}} \sim f_{\mathrm{t}}$, can be easily done by multiplying each set of quantiles $(h'_{l-1} < u_{n,l} \leq h'_l)$ from u_{n}, with σ_l and then by adding μ_l (where $h'_{l-1} = \frac{h_l - 1 - \mu_l}{\sigma_l}$, $h'_l = \frac{h_l - \mu_l}{\sigma_l}$ and $u_{n,l} = \frac{u_l - \mu_l}{\sigma_l}$). This scheme can be easily applied to any type of empirical pdf, however in cases where the empirical pdf highly deviates from a Normal pdf, a large number of segments may be acquired and the process' pdf be poorly interpreted.

Here, we observe that the left and right branch of the averaged empirical pdf can be very well approximate by two Gaussian distributions. Thus, we approximate the pdf of the process with 2 segments (o = 2), with parameters shown in Fig. 5(b), with Pearson correlation coefficient R^2 = 0.995, between the empirical and the modelled pdf of equation (45):

$$f_{\mathrm{t}}(u) = \begin{cases} \mathrm{N}(0,1), & -\infty < u \leq 1 \\ \mathrm{N}(-0.3,1.3), & 1 < u < \infty \end{cases} \tag{45}$$

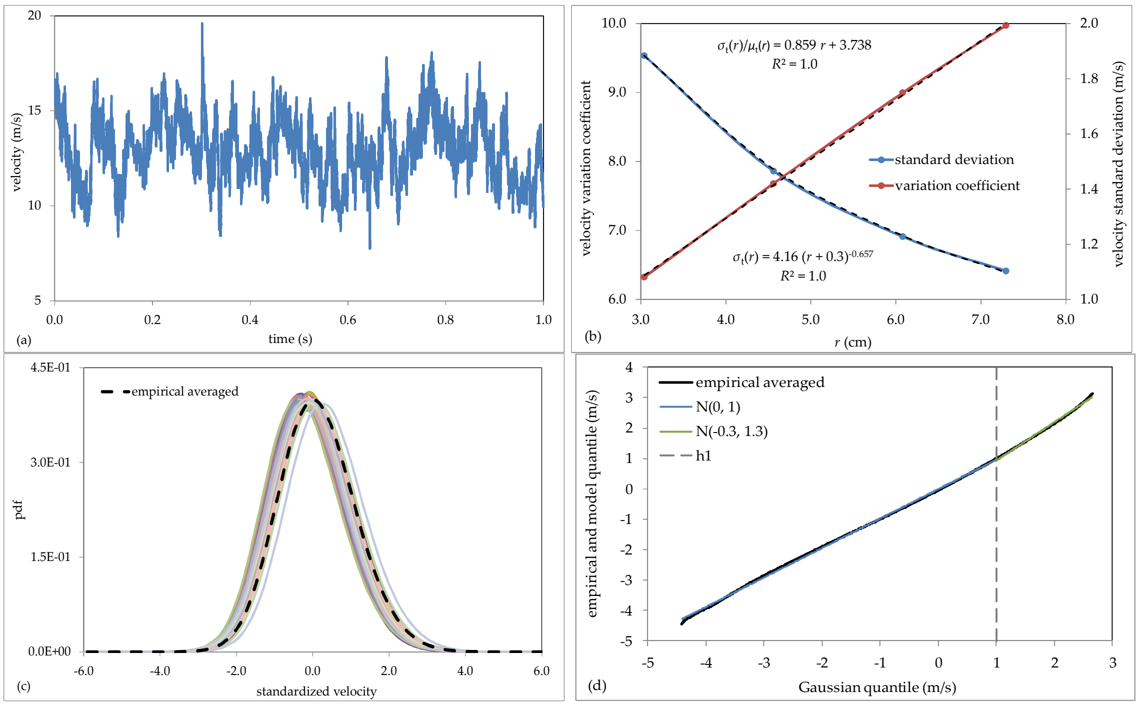

Fig. 5: Data preliminary analysis: (a) 1 s time window of one of the raw time-series; (b) averaged velocity mean $\mu_{\mathrm{t}}(r)$ divided by the averaged velocity standard deviation $\sigma_{\mathrm{t}}(r)$ (variation coefficient) and averaged velocity standard deviation $\sigma_{\mathrm{t}}(r)$ as a function of r, along the longitudinal axis, as well as their fitted curves (black dashed lines); (c) empirical pdf's of the standardized time-series (multi-coloured lines) by subtracting $\mu_{\mathrm{t}}(r)$ and dividing with $\sigma_{\mathrm{t}}(r)$ each time-series and the empirical averaged pdf; (d) qq-plot of averaged empirical pdf *vs* standard Gaussian pdf, i.e. N(0,1), along with modelled pdf from equation 45 (all parameters in m/s).

In Fig. 6, we show the climacograms, autocovariances, variograms, power spectra, CBV's and CBS's from all 40 standardized time-series, their averaged values and the corresponding values of the 38th time-series. Assuming that these averaged values are near the process' expected ones, we can fit a stochastic model based on all the stochastic tools examined, and particularly the ones with the smallest statistical error for each scale, lag and frequency. We observe (Fig. 6g-h) that the large scale autocovariance and climacogram expected LLD's are both larger than -1 and that the power spectrum and CBS low frequency expected LLD's are larger than 0. Hence, it is most probable that the process exhibits HK behaviour.

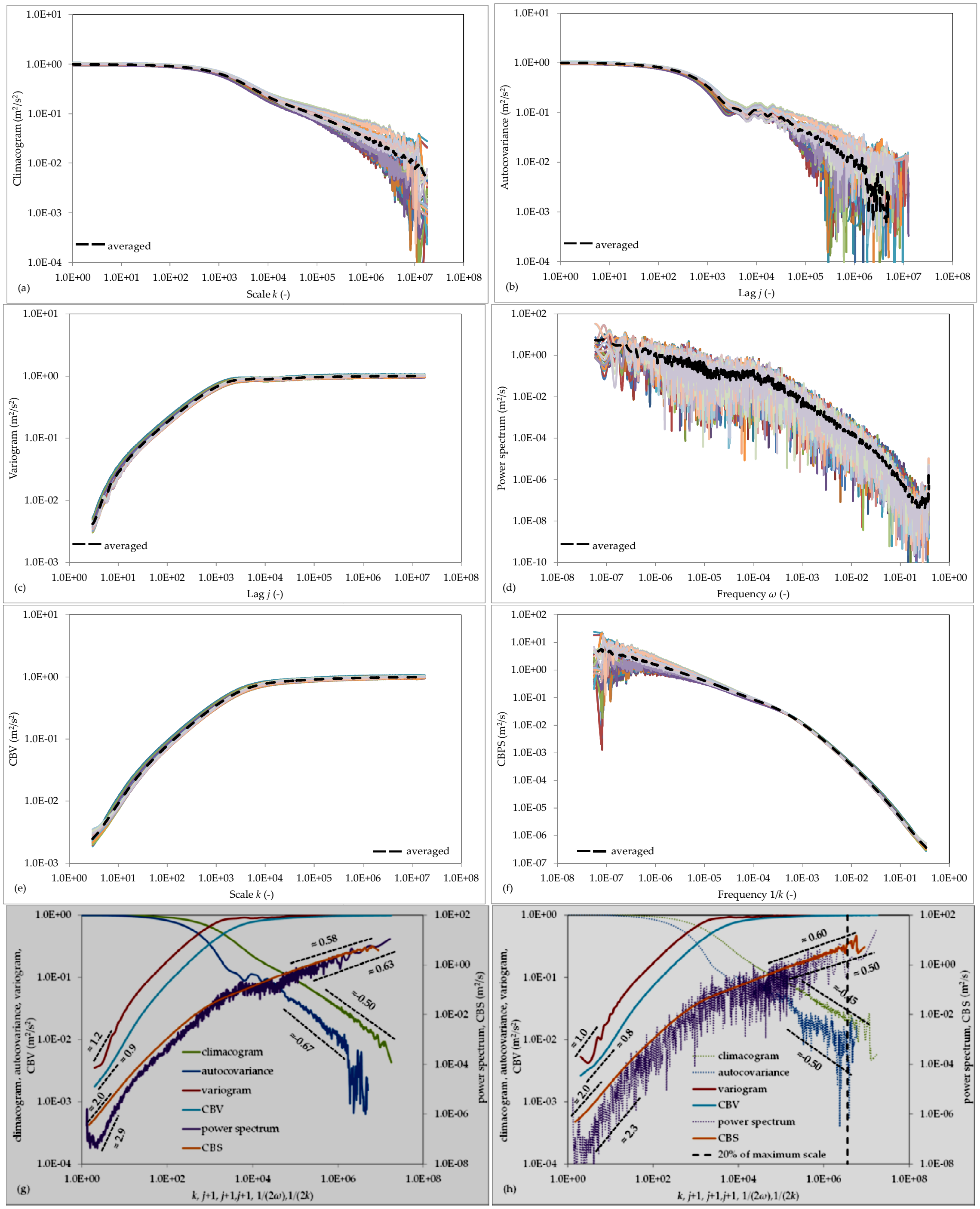

Fig. 6: Data analysis: (a) climacograms, (b) autocovariances, (c) variograms, (d) power spectra, (e) CBV and (f) CBS (with $\gamma(0)$ taken from the model in Table 2), of all the 40 time-series (multi-coloured lines) as well as their averaged values (black dashed lines); (g) all averaged values along with their averaged LLD's at large scales, lags and inverse frequencies and (h) those of the 38th time-series. Note that we use scales, lags and inverse frequencies up to the 20% of the maximum scale for our calculations, following the rule of thumb proposed in Koutsoyiannis (2003), Dimitriadis and Koutsoyiannis (2015).

Primarily, we try to best fit the climacogram-based stochastic tools (for reasons that will be explained later) and secondarily, the variogram for the intermediate lags (see in Fig. 7). To estimate the process parameters in Table 2, a dimensionless fitting error is considered (as in Dimitriadis and Koutsoyiannis 2015):

$$\mathrm{FE}_{\theta} = \sum_z \left(\frac{\mathrm{E}[\underline{\hat{\theta}}(z)] - \hat{\underline{\theta}}_d^{(\Delta)}(z)}{\mathrm{E}[\underline{\hat{\theta}}(z)]}\right)^2 \tag{46}$$

where $\hat{\underline{\theta}}_d^{(\Delta)}$ is the empirical stochastic tool estimated from the data, $\mathrm{E}[\underline{\hat{\theta}}]$ the expected one estimated from the model and z the corresponding to the stochastic tool scale, lag or frequency.

The optimization analysis results in scale parameters $\lambda_1 = 0.422$ m²/s² and $\lambda_2 = 0.592$ m²/s², shape parameters $q_1 = 19.6$ ms and $q_2 = 1.45$ ms, fractal parameter $a = 1.4$ and HK parameter $b = 0.32$, with correlation coefficient R^2 approximating 1.0 for the climacogram and CBV, 0.99 for the CBS and variogram, 0.95 for the autocovariance and 0.8 for the power spectrum.

Applying the L'Hôpital's rule and through mathematical calculations, we find that the fractal dimension of the process in Table 2 is affected only by the exponent α of the powered-exponential process and the Hurst coefficient only by the exponent b of the gHK one. Thus, process' fractal dimension and Hurst coefficient are estimated (based on the definition in equation (11) and (12) and Gneiting and Schlather 2004, analysis) as:

$$F = 2 - \frac{\alpha}{2} = 1.3 \tag{47}$$

$$H = 1 - \frac{b}{2} = 0.84 \tag{48}$$

Finally, to test the validity of our initial assumption, that for the specific model in Table 2 and the estimated parameters the classical estimators of the climacogram-based stochastic tools have the smallest error ε if compared to the autocovariance, variogram and power spectrum ones, we proceed as follows. We calculate the statistical error for each stochastic tool via Monte Carlo analysis (since we lack analytical expressions for the variance of the expected values):

$$\varepsilon_{\theta} = \frac{\mathrm{E}\left[\left(\underline{\hat{\theta}} - \theta\right)^2\right]}{\theta^2} = \varepsilon_{\theta,v} + \varepsilon_{\theta,b} \tag{49}$$

where we have decomposed the dimensionless mean square error into a variance and a bias term (see in Dimitriadis and Koutsoyiannis 2015),

$$\varepsilon_{\theta,v} = \mathrm{var}[\underline{\hat{\theta}}]/\theta^2 \tag{50}$$

$$\varepsilon_{\theta,b} = \left(\theta - \mathrm{E}[\underline{\hat{\theta}}]\right)^2/\theta^2 \tag{51}$$

where θ is the examined stochastic tool, $\varepsilon_{\theta,b}$ can be easily estimated from equations in Tables A1-A6 and $\varepsilon_{\theta,v}$ is calculated from the Monte Carlo analysis since we lack analytical expressions.

Thus, we produce 40 time-series with $n = 36\times10^6$ using the SMA algorithm (Koutsoyiannis 2000 and 2015), which can replicate any stochastic process. Then, we compare the errors ε for each stochastic tool for 81 points logarithmically distributed from 1 to n (Fig. 8). Note that in Fig. 8, we try to show all estimates within a single plot for comparison. The inverse frequency in the horizontal axis is set to $1/(2\omega)$, in order to vary between 1 and $n/2$, and the lag to j+1, so as the estimation of variance at $j = 0$ is also shown in the log-log plot. From the results of this analysis, it can be observed that the initial choice of the climacogram-based stochastic tools (and the variogram's for a small window of intermediate LLD's) to interpret the empirical process, is proven valid for the current model structure, model parameters and examined range of scales, with the power spectrum exhibiting the largest errors.

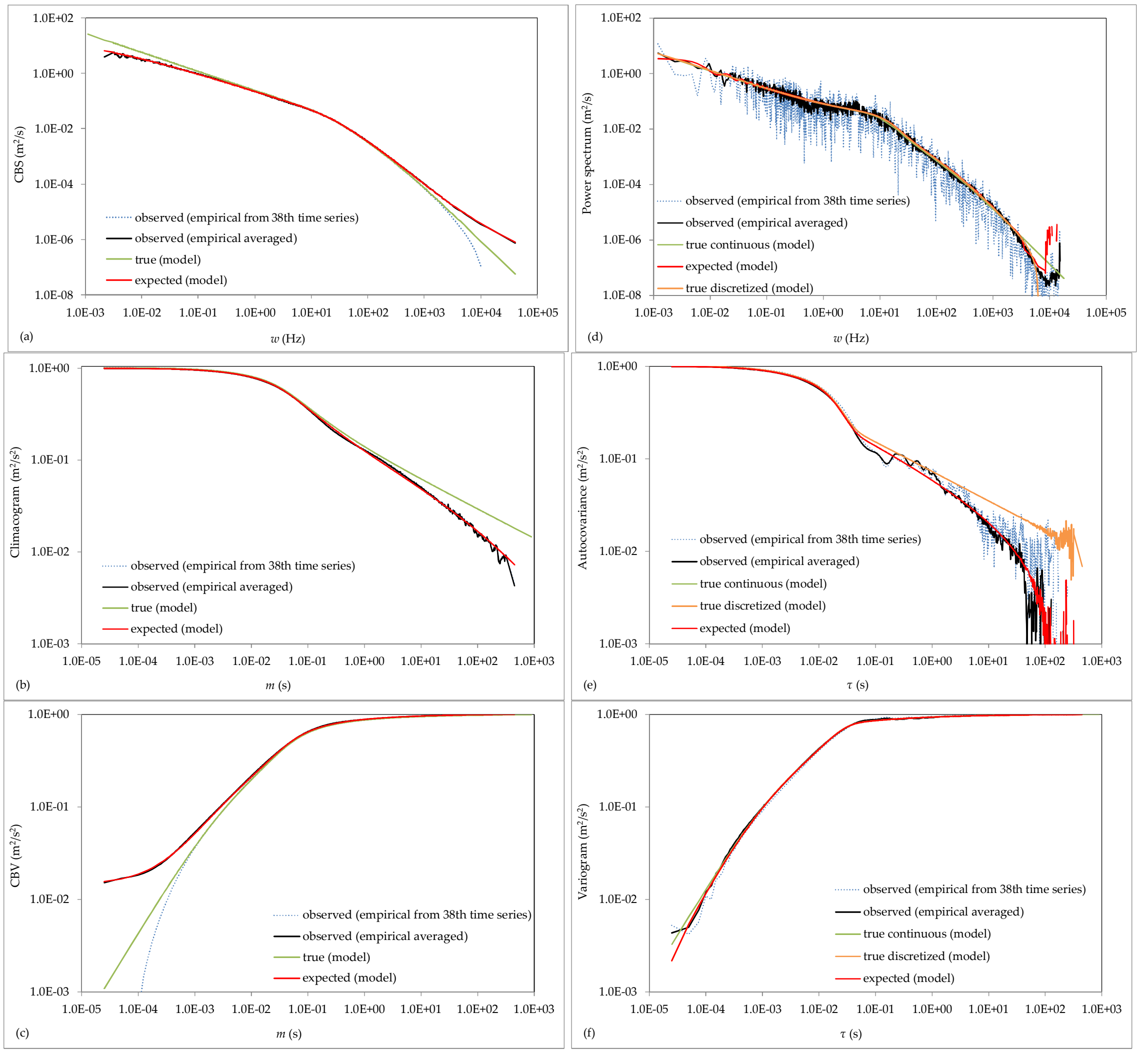

Fig. 7: Stochastic modelling: true in continuous time (estimated from the model), true in discrete time (estimated from the model), expected (estimated from the model), empirical averaged (estimated from all 40 time-series) and observed (estimated from the 38th time-series), for the (a) CBS, (b) climacogram, (c) CBV, (d) power spectrum, (e) autocovariance and (f) variogram.

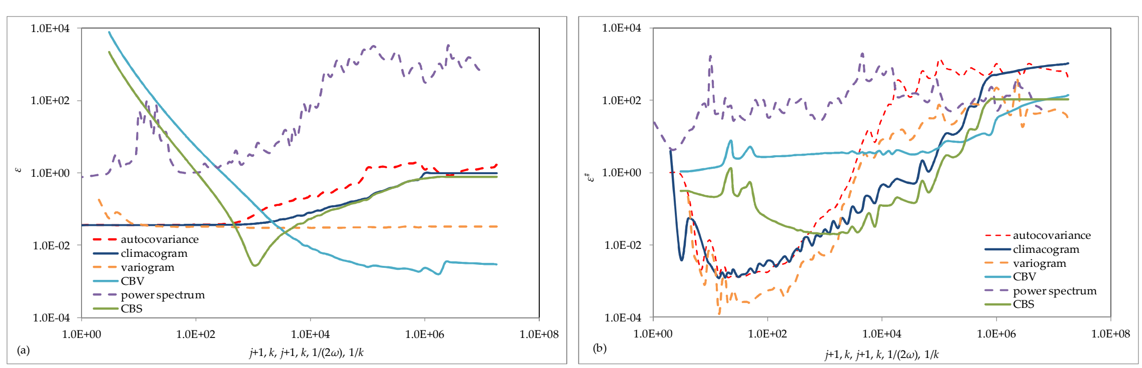

Fig. 8: Dimensionless errors (a) ε_θ and (b) $\varepsilon_{\theta^\#}$ of the climacogram, autocovariance, variogram, CBV, power spectrum and CBS calculated from 40 synthetic series with n = 36×10⁶, based on the process in Table 2. Note that the LLD's included in $\varepsilon_{\theta^\#}$ estimations are calculated using equation (10).

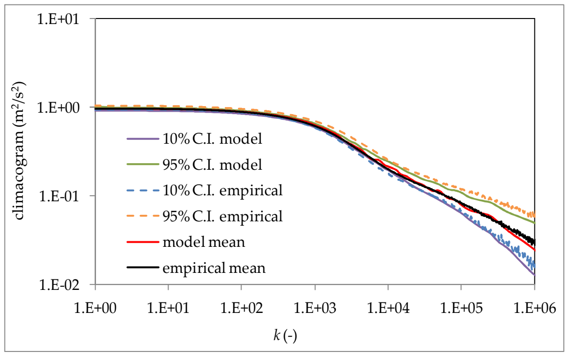

Fig. 9: Empirical *vs* modeled 10% and 95% confidence intervals based on the climacogram (approximately up to the 20% of maximum scale).

Additionally, we estimate the empirical process low and high confidence intervals (for the climacogram only) for the chosen model and fitted parameters around 10% and 95%, respectively (Fig. 9). Note that the reason we apply the model to the expected value of the empirical process and not to the mode is because it is much simpler due to the existence of analytical expressions of the expected values. The method of maximum likelihood is far too complicated and time-consuming (due to the lack of analytical expressions) but it offers better interpretation of the process. However, in cases where there are multiple realizations of the process (as in the current application so that we can have an estimate of the expectation of the process), the proposed in this paper method combines both simplicity and ample statistical basis.

4.2 Application to atmospheric wind speed

In this section we show the stochastic analysis of a time-series of one month (Fig. 10), consisted of high resolution ($\Delta \approx D = 0.1$ s) atmospheric longitudinal wind speed (measured in m/s). This is recorded by a sonic anemometer on a meteorological tower, located at Beaumont KS and are provided by NCAR/EOL (http://data.eol.ucar.edu/). First, we divide the time-series into 3 sets, each of which includes around 1400 time-series of 10 min duration and with marginal empirical variances 0.15, 0.5 and 1.4 m^2/s^2, respectively (Fig. 11). We have chosen this process since it is of high importance in hydrometeorology and it includes a large variety of marginal variances. In Fig. 11, one may clearly observe the transition from a process with low marginal variance having a power spectrum with a drop in the intermediate scales (like in the grid-turbulence application), to the one with larger marginal variance power spectrum (with no drop). This again shows the importance of the type of model we propose in this paper (Table 2), which can describe a great variety of natural processes' behaviours.

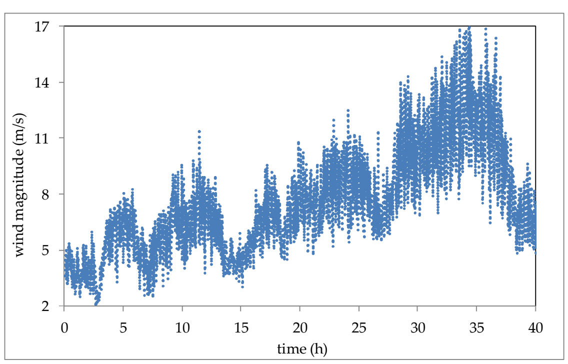

Fig. 10: Part of the wind speed time-series provided by NCAR/EOL (http://data.eol.ucar.edu/).

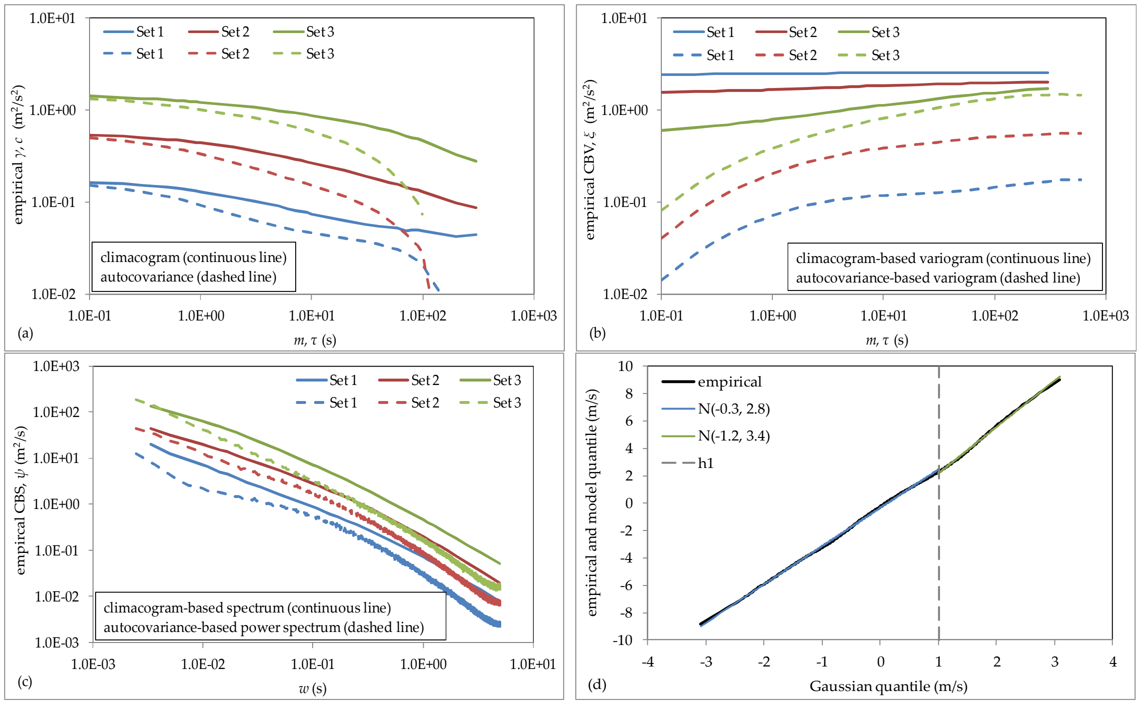

Fig. 11: Averaged empirical (a) climacograms and autocovariances, (b) CBV and variograms, (c) CBS and power spectra (for the three sets) and (d) qq-plot of empirical pdf *vs* standard Gaussian pdf (for the original time-series), along with modelled pdf from equation 42 (all parameters in m/s).

However, it would be more appropriate to apply separately first, the powered-exponential, gHK and Gneiting model (see equation 52), if the empirical process seems to have two distinctive areas (like the 2nd and 3rd set of wind speed). In the next equations, we present stochastic tools for the Gneiting process, with some alterations to include cases of $H \rightarrow 0$ and white noise behaviour, i.e. $H = 0.5$ (so as to be also consistent with the HK process, cf. Koutsoyiannis 2015):

$$c(\tau) = \frac{\lambda(1-b)(2-b)}{(1+(|\tau|/q)^{a})^{b/a}} \tag{52}$$

$$\gamma(m) = \lambda\left(2\ {}_1\mathrm{F}_2\left[\frac{1}{a}, \frac{b}{a}, 1+\frac{1}{a}, -\left(\frac{m}{q}\right)^{a}\right] - {}_1\mathrm{F}_2\left[\frac{2}{a}, \frac{b}{a}, \frac{2+a}{a}, -\left(\frac{m}{q}\right)^{a}\right]\right) \tag{53}$$

with $a, b \geq 0$ and $\lambda(1-b)(2-b)$ the process' variance (the expressions for the rest tools can be found in Appendix and cannot be written in an analytical form).

Applying the same methodology as in the previous section, the optimization analysis (from the best fitted model of Table 2) results for the 1st set in scale parameters: $\lambda_1 = 0.115$ m²/s² and $\lambda_2 = 2.502$ m²/s², shape parameters $q_1 = 0.484$ s and $q_2 = 103.7$ s, fractal parameter $a = 0.6$ ($F = 1.7$) and HK parameter $b = 0.02$ ($H = 0.99$). For the 2nd set, the best fit corresponds to the Gneiting process (equation 52): $\lambda = 1.124$ m²/s², $q = 0.029$ s, $a = 2$ ($F = 1$) and $b = 0.04$ ($H = 0.98$). Finally, for the 3rd set, the best fit corresponds to the gHK process with parameters: $\lambda_2 \approx 6$ m²/s², $q_2 \approx 0.4$ s and $b \approx 0.04$ ($H = 0.98$). The fitted model (in terms of the climacogram) can be viewed in Fig. 12.

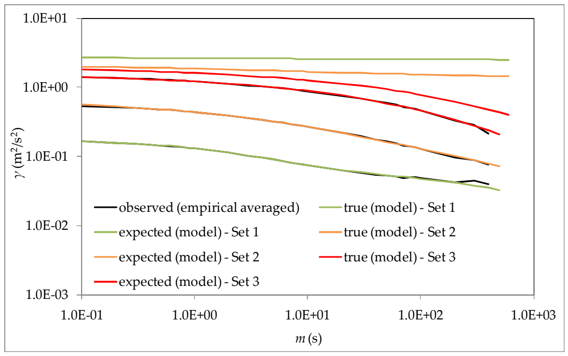

Fig. 12: True, expected and empirical (averaged) climacogram values for the wind process stochastic simulation.

4.3 Application to high resolution precipitation

In this section we show the stochastic analysis of three time-series (Fig. 13) with high resolution ($\Delta \approx D$ = 10 s) precipitation intensities (measured in mm/h). These episodes are recorded during various weather states (high and low rainfall rates) and provided by the Hydrometeorology Laboratory at the Iowa University (for more information concerning these episodes and various stochastic analyses, see Georgakakos et al., 1994; Papalexiou et al. 2011; Koutsoyiannis and Langousis 2011 ch. 1.5).

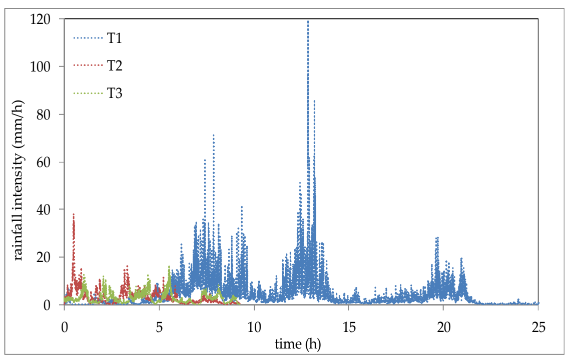

Fig. 13: Three precipitation episodes provided by the Hydrometeorology Laboratory at the Iowa University (see Georgakakos et al. 1994).

In this case, we treat each episode separately and so, we fit the expected value of the model to the empirical process (a more statistically correct way would be to work with the mode). Note that the normalization scheme proposed in this paper would require around five Gaussian functions (due to the highly skewed probability function) and so, we should use a simpler scheme (e.g. Papalexiou et al. 2011). Applying the same methodology for the stochastic simulation as in the previous sections, the optimization analysis for T1 results to the model in Table 2, with: $\lambda_1 = 18.0$ mm²/h² and $\lambda_2 = 110.0$ mm²/h², shape parameters $q_1 = 18.47$ s and $q_2 = 4250.0$ s, fractal parameter $a = 1.44$ ($F = 1.28$) and HK parameter $b = 0.12$ ($H = 0.94$). For the T2, the best fit corresponds to the Gneiting process (equation 52): $\lambda = 20.153$ mm²/h², $q = 33.016$ s, $a = 1.94$ ($F \approx 1$) and $b = 0.09$ ($H \approx 0.95$). Finally, for T3 the best fit corresponds to the gHK process in Table 2, with parameters: $\lambda_1 = 13.2$ mm²/h², shape parameters $q_1 = 111.7$ s and HK parameter $b = 0.13$ ($H \approx 0.93$).

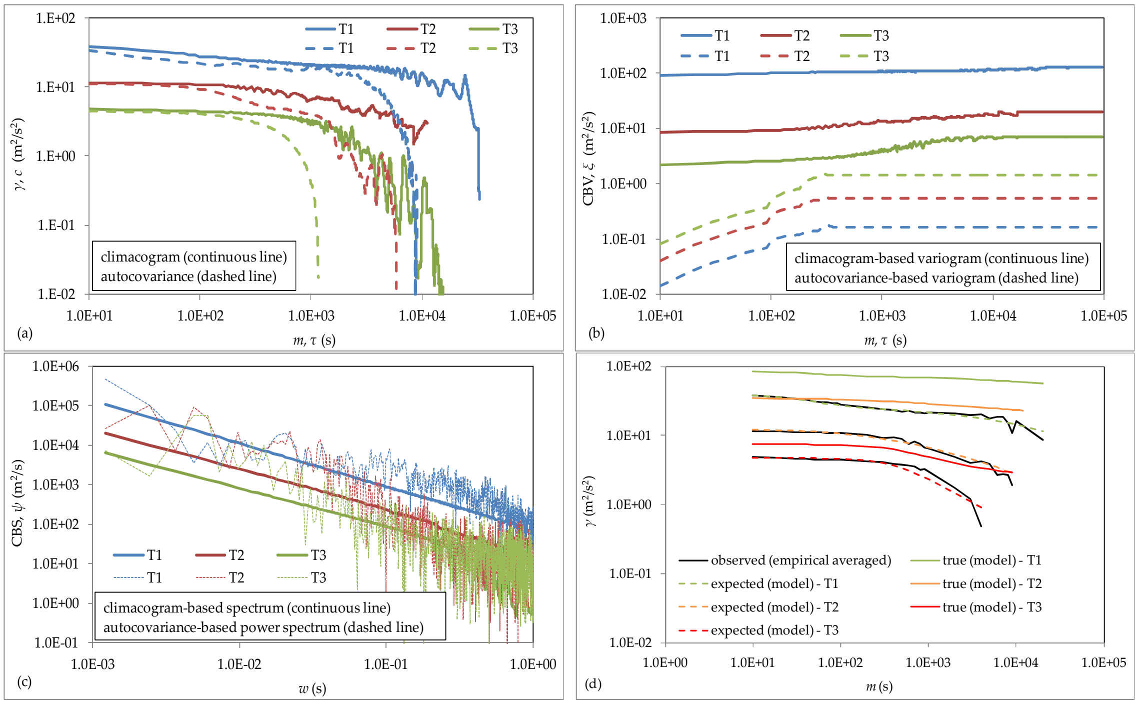

Fig. 14: Averaged empirical (a) climacograms and autocovariances, (b) CBV and variograms, (c) CBS and power spectra for T1, T2 and T3, and (d) true, expected and empirical (averaged) climacogram values for the rainfall processes stochastic simulation.

5. Summary and conclusions

Studying turbulence is very helpful in hydrology, as it can provide us with long time-series, enabling us to focus on the crucial, for hydrological processes, long term properties. Also, it is important in the interpretation of hydrological (macroscale) processes as turbulence generates and drives most of them through microscale mechanisms. In this paper, we investigate the most common power-spectrum based stochastic models of stationary and isotropic turbulence. We see that these models have a high order of complexity when they are multiplied with each other in order to be combined into a single equation. Also, most of these models lack stochastic interpretation (as they cannot easily be analyzed into basic stochastic processes such as powered-exponential or power-type decay of autocovariance with lag). Moreover, we remark that these models can lead to natural process misinterpretation due to the power spectrum identical asymptotic power spectrum behaviours for stochastically different geophysical processes, e.g. Markovian and gHK with b=1. Finally, these models do not include important stochastic parameters, such as Hurst coefficient and fractal dimension, thus it often results in violating basic stochastic asymptotic properties such as theoretical limits of the Hurst coefficient, in case that Hurst-Kolmogorov (HK) behaviour is observed.

Using the stochastic framework shown in Appendix, we propose a more simple, flexible and robust model in Table 2 that can incorporate both powered-exponential and HK behaviours in a wide range of scales. This model also exhibits the Kolmogorov's log-log derivative of '-5/3' in the intermediate frequencies without assuming intermediate power law functions. Furthermore, it gives a possible explanation of the high frequency spike frequently met in power spectra of turbulence time-series that is probably caused by the process discretization and bias. This model is also tested with high resolution grid (nearly-isotropic) turbulence velocity measurements of laboratory scale, exhibiting an excellent agreement. Additionally, we show two examples of hydrometeorological processes (including wind speed and precipitation time-series), which often present similar behaviours to the microscale of turbulence. Moreover, we highlight the advantages of using more than one stochastic tools to interpret the natural process based on the ones with smaller uncertainty and statistical errors. More specifically, we compare the climacogram with the autocovariance, the climacogram-based variogram with the classical autocovariance-based variogram and the climacogram-based spectrum with the classical power spectrum. We find that combining together climacogram-based stochastic tools results in smaller uncertainty and statistical errors in regular and log-log derivatives over the longest range of scales, lags and frequencies, with the power spectrum giving the largest errors. Finally, we estimate the two parameters characterizing the self-similarity of the examples of turbulence, wind speed and precipitation processes, namely the fractal dimension and Hurst coefficient, which refer to small and large time scales respectively.

Acknowledgements

We would like to thank the Associate Editor Elena Volpi as well as the Reviewers Francesco Laio and Federico Lombardo for their instructive and helpful comments, including some important limitations on power-spectrum's smoothing techniques.

Funding

This work was partly funded by the Greek General Secretariat for Research and Technology through the research project Combined REnewable Systems for Sustainable Energy DevelOpment (CRESSENDO, grant number 5145).

References

Avila, M., Willis, A.P., and Hof, B., 2010. On the transient nature of localized pipe flow turbulence. *Journal of Fluid Mechanics*, 646, 127-136.

Batchelor, G.K., 1959. *The Theory of Homogeneous Turbulence*. Cambridge University Press.

Batchelor, G.K. & Townsend, A.A. 1949. The nature of turbulent motion at large wave-numbers. *Proceedings of Royal Society of London A*, 199, 238–255.

Beran, J., Feng, Y. and Ghosh, S., 2013. Long-Memory Processes, Probabilistic Properties and Statistical Methods. *Springer*.

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

Chamorro, P.L. and Porte-Agel, F., 2009. A Wind-Tunnel Investigation of Wind-Turbine Wakes: Boundary-Layer Turbulence Effects. *Boundary-Layer Meteorology*, 132: 129-149.

Charakopoulos, A.K., Karakasidis, T.E., Papanicolaou, P.N. and Liakopoulos, A., 2014a. Nonlinear time-series analysis and clustering for jet axis identification in vertical turbulent heated jets. *Physical Review E*, 89, 032913.

Charakopoulos, A.K., Karakasidis, T.E., Papanicolaou, P.N. and Liakopoulos, A., 2014b. The application of complex network time-series analysis in turbulent heated jets. *CHAOS*, 24, 024408.

Davidson, P.A., 2000. Was Loitsyansky correct? A review of the arguments. *Journal of Turbulence*,1,N6.

Dimitriadis, P., and Papanicolaou, P., 2012. Statistical analysis of turbulent positively buoyant jets. *European Geosciences Union General Assembly 2012*, Geophysical Research Abstracts, Vol. 14, Vienna, EGU2012-12672, European Geosciences Union.

Dimitriadis, P., and Koutsoyiannis, D., 2015. Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes. *Stochastic Environmental Research & Risk Assessment*, doi: 10.1007/s00477-015-1023-7.

Faisst, H., and Eckhardt, B., 2004. Sensitive dependence on initial conditions in transition to turbulence in pipe flow. *Journal of Fluid Mechanics*, 504, 343-352.

Falkovich, G., Fouxon, A. and Stepanov, M.G., 2002. Acceleration of rain initiation by cloud turbulence. *Nature*, (419) 151–154.

Frisch, U., 2006. *The Legacy of A. N. Kolmogorov*. Cambridge University Press.

Georgakakos, K.P., Bae, D.H. and Cayan, D.R. 1994. Hydro-climatology of continental watersheds, 1, Temporal analyses. *Water Resources Research*, 31(3), 655–675.

Gneiting, T. 2000. Power-law correlations, related models for long-range dependence, and their simulation. *Journal of Applied Probability* 37, 1104–1109.

Gneiting, T. and Schlather, M., 2004. Stochastic models that separate fractal dimension and the Hurst effect. *SIAM review*, 46(2),269-282.

Gneiting T., Sevcikova, H. and Percival, D.B., 2012. Estimators of Fractal Dimension: Assessing the Roughness of Time-series and Spatial Data. *Statistical Science*, 27 (2): 247-277.

Goldstein, M.L. and Roberts, D.A., 1995. Magnetohydrodynamic turbulence in the solar wind. *Annual Review of Astronomics and Astrophysics*, 33:283-325.

Heisenberg, W., 1948. On the theory of statistical and isotropic turbulence. *Proceedings of the Royal Society of London,* A(195) 402.

Helland K.N. and Van Atta, C.W., 1978. The 'Hurst phenomenon' in grid turbulence. *Journal of Fluid Mechanics*, 85(3): 573-589.

Hurst, H.E., 1951. Long term storage capacities of reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 776–808.

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

Kolmogorov, A.N., 1940. Wiener's spiral and some other interesting curves in Hilbert space. *Doklady Akademii Nauk SSSR*, 26.

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

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

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

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

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

Koutsoyiannis, D., 2011a. Hurst-Kolmogorov dynamics and uncertainty, *Journal of the American Water Resources Association*. 47 (3), 481–495.

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

Koutsoyiannis, D., 2013a. *Encolpion of stochastics: Fundamentals of stochastic processes* [online]. Department of Water Resources and Environmental Engineering, National Technical University of Athens Available from: http://www.itia.ntua.gr/en/docinfo/1317. [Accessed 4 November 2004].

Koutsoyiannis, D., 2013b. Hydrology and Change. *Hydrological Sciences Journal*, 58 (6), 1177–1197.

Koutsoyiannis, D., 2014. Reconciling hydrology with engineering. *Hydrology Research*, 45 (1), 2–22.

Koutsoyiannis, D., 2015. Generic and parsimonious stochastic modelling for hydrology and beyond. *Hydrological Sciences Journal*, doi:10.1080/02626667.2015.1016950.

Koutsoyiannis, D., and Langousis, A. 2011. Precipitation, Treatise on Water Science, edited by P. Wilderer and S. Uhlenbrook, 2, 27–78, Academic Press, Oxford.

Kraichnan, R.H., 1959. The structure of isotropic turbulence at very high Reynolds number. *Journal of Fluid Mechanics*, (5) 497.

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

Kuik, D. J., Poelma, C., and Westerweel, J., 2010. Quantitative measurement of the lifetime of localized turbulence in pipe flow. *Journal of Fluid Mechanics*, 645, 529-539.

Lombardo, F., Volpi, E. and Koutsoyiannis, D., 2013. Effect of time discretization and finite record length on continuous-time stochastic properties. *IAHS - IAPSO - IASPEI Joint Assembly*, Gothenburg, Sweden, International Association of Hydrological Sciences, International Association for the Physical Sciences of the Oceans, International Association of Seismology and Physics of the Earth's Interior (http://itia.ntua.gr/1380).

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

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

Richter, J.P., 1939 (first published in 1883). *The Literary Works of Leonardo da Vinci*. 2nd ed. London and New York: Oxford University Press.

Mandelbrot, B.B., and Wallis, J.R., 1968. Noah, Joseph, and operational hydrology. *Water Resources Research*, 4 (5), 909-918.

McDonough, J.M., 2007. Introductory Lectures on Turbulence, Physics, Mathematics and Modeling , University of Kentucky.

Nordin C.F., McQuivey, R.S. and Mejia, J.M., 1972. Hurst phenomenon in Turbulence. *Water Resources Research*, 8 (6).

Papalexiou, S.M., Koutsoyiannis, D., and Montanari, A., 2011. Can a simple stochastic model generate rich patterns of rainfall events? *Journal of Hydrology*, 411 (3-4), 279–289.

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

Richardson, L.F., 1922. *Weather Prediction by Numerical Process*. Cambridge University Press.

Rinaldo, A., 2006. Introduction to special issue on Rain, Rivers, and Turbulence: A view from hydrology. *Water Resources Research*, 42, W06D01.

Saffman, P.G., 1967. The large-scale structure of homogeneous turbulence. *Journal of Fluid Mechanics*, 27 (3), 581-593.

Stoica, P. And Moses, R., 2004. *Spectral analysis of signals*, Prentice Hall, Upper Saddle River, New Jersey.

Taylor, G.I., 1935. Statistical Theory of Turbulence I-IV. *Proceedings of the Royal Society of London,* A (151), 421-454.

Tessarotto, M. and Asci, C., 2010. On the behaviour of homogeneous, isotropic and stationary turbulence. *Fluid Dynamics,* arXiv:1003.1475 .

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

Wackernagel, H., 1995. Multivariate Geostatistics: an Introduction with Applications. 1st ed. Berlin: Springer.

Wright, J.R. and Cooper, J.E. 2007. Introduction to Aircraft Aeroelasticity and Loads. *John Wiley & Sons Ltd*.

Yaglom, A.M., 1987. Correlation Theory of Stationary and Related Random Functions, Vol. I: Basic Results, New York: Springer.

Appendix

Here, we present a climacogram-based stochastic framework (Koutsoyiannis 2013a; Dimitriadis and Koutsoyiannis 2015). All observed time-series are subject to a sampling time interval D, often fixed by the observer and a response time Δ ($\leq D$) of the instrument (Fig. A1), that both affect the estimation of the statistical properties of the continuous time process $\underline{x}(t)$. Thus, the discrete time stochastic process $\underline{x}_i^{(\Delta)}$, can be calculated from $\underline{x}(t)$ as:

$$\underline{x}_i^{(\Delta,D)} = \frac{\int_{(i-1)D}^{(i-1)D+\Delta} \underline{x}(\xi)\mathrm{d}\xi}{\Delta} \tag{A1}$$

where $i \in [1, n]$ is an index representing discrete time, $n = \lfloor T/\Delta \rfloor$ is the total number of observations and $T \in [0, \infty)$ is the time length of observations.

For simplicity reasons here, we assume that $D \approx \Delta > 0$, which is also practical for samples with small D (as the one shown in the application in Section 4). An example of the Markovian process with $D \neq \Delta$ can be found in Dimitriadis and Koutsoyiannis (2015). Additional examples and stochastic tools for the two special cases $D = \Delta > 0$ and $D > \Delta = 0$, can be found in Koutsoyiannis (2013a). From these analyses, one can conclude that the differences between the two extreme cases are often small for small D.

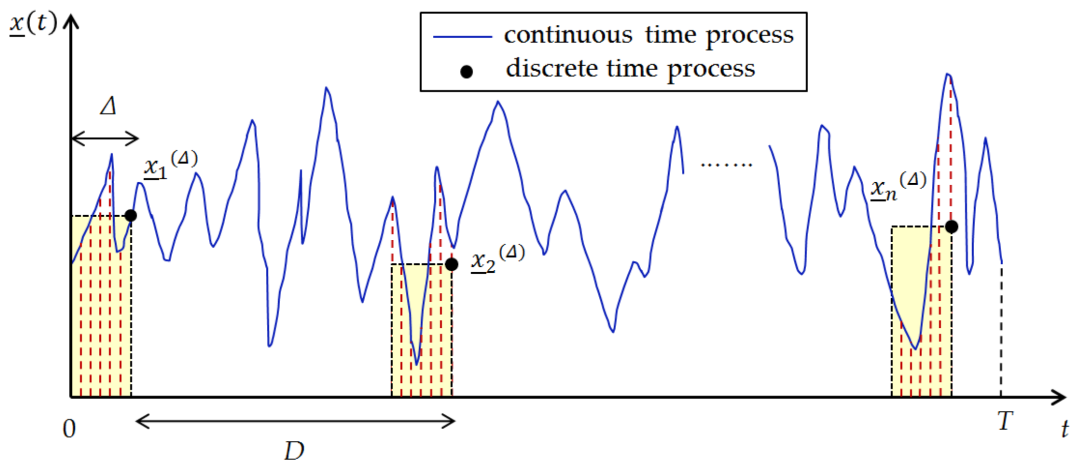

Fig. A1: An example of a continuous time process sampled at time intervals D for a total period T and with instrument response time Δ.

In Table A1, we introduce the climacogram definition in case of a stochastic process in continuous time (equation A2) and in discrete time (equation A3), a widely used climacogram estimator (equation A4) and climacogram estimation (based on the latter estimator) and expressed in function with the true climacogram (equation A5). In Tables A2 and A3, we introduce the CBV as well as the CBPS.

Moreover, in Table A4, we define the autocovariance function in case of a stochastic process in discrete time (equation A15), a widely used autocovariance function estimator (equation A16) as well as an estimation based on the latter estimator and expressed in function with the true climacogram (equation A17, derived in Dimitriadis and Koutsoyiannis 2015). In Tables A5 and A6, we define the autocovariance-based classical variogram and power spectrum.

Table A1: Climacogram definition and expressions for a process in continuous and discrete time, along with the properties of its estimator.

Type	Climacogram	
continuous	$\gamma(m) := \mathrm{Var}\left[\int_0^m \underline{x}(\xi)\mathrm{d}\xi\right]/m^2$ where $m \in \mathbb{R}^+$	(A2)
discrete	$\gamma_\mathrm{d}^{(\Delta)}(k) := \dfrac{\mathrm{Var}\left[\sum_{l=1}^{k} \underline{x}_l^{(\Delta,D)}\right]}{k^2} = \gamma(k\Delta)$ where $k \in \mathbb{N}$ is the dimensionless scale for a discrete time process	(A3)
classical estimator	$\underline{\hat{\gamma}}_\mathrm{d}^{(\Delta)}(k) = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(\dfrac{1}{k}\left(\sum_{l=k(i-1)+1}^{ki} \underline{x}_l^{(\Delta)}\right) - \dfrac{\sum_{l=1}^{n} \underline{x}_l^{(\Delta)}}{n}\right)^2$	(A4)
expectation of classical estimator	$\mathrm{E}\left[\underline{\hat{\gamma}}_\mathrm{d}^{(\Delta)}(k)\right] = \dfrac{1-\gamma(n\Delta)/\gamma(k\Delta)}{1-k/n}\gamma(k\Delta)$	(A5)

Table A2: Climacogram-based variogram definition and expressions for a process in continuous and discrete time, along with the properties of its estimator.

Type	Climacogram-based variogram	
continuous	$\xi(m) := \gamma(0) - \gamma(m)$	(A6)
discrete	$\xi_\mathrm{d}^{(\Delta)}(k) := \gamma(0) - \gamma(k\Delta)$	(A7)
classical estimator	$\underline{\hat{\xi}}_\mathrm{d}^{(\Delta)}(k) = \gamma(0) - \underline{\hat{\xi}}_\mathrm{d}^{(\Delta)}(k)$	(A8)
expectation of classical estimator	$\mathrm{E}\left[\underline{\hat{\xi}}_\mathrm{d}^{(\Delta)}(k)\right] = \gamma(0) - \mathrm{E}\left[\underline{\hat{\xi}}_\mathrm{d}^{(\Delta)}(k)\right]$	(A9)

Table A3: Climacogram-based spectrum (pseudospectrum) definition and expressions for a process in continuous and discrete time, along with the properties of its estimator.

Type	Climacogram-based spectrum	
continuous	$\psi(m) := \dfrac{2\gamma(1/w)}{w}\left(1 - \dfrac{\gamma(1/w)}{\gamma(0)}\right)$ where $w \in \mathbb{R}$ is the frequency for a continuous time process (in inverse time units) and is equal to w=1/m.	(A10)
discrete	$\psi_{\mathrm{d}}^{(\Delta)}(\omega) := \dfrac{2\gamma(1/\omega)}{\omega}\left(1 - \dfrac{\gamma(1/\omega)}{\gamma(0)}\right)$ where $\omega \in \mathbb{R}$ is the frequency for a discrete time process (dimensionless; $\omega = w\Delta$)	(A11)
classical estimator	$\hat{\underline{\psi}}_{\mathrm{d}}^{(\Delta)}(\omega) = \dfrac{2\gamma(1/\omega)}{\omega}\left(1 - \dfrac{\gamma(1/\omega)}{\gamma(0)}\right)$	(A12)
expectation of classical estimator	$\mathrm{E}\left[\hat{\underline{\psi}}_{\mathrm{d}}^{(\Delta)}(\omega)\right] = \dfrac{2\mathrm{E}[\gamma(1/\omega)]}{\omega}\left(1 - \dfrac{\mathrm{E}[\gamma(1/\omega)]}{\gamma(0)} - \dfrac{\mathrm{Var}[\gamma(1/\omega)]}{\gamma(0)\mathrm{E}[\gamma(1/\omega)]}\right)$	(A13)

Table A4: Autocovariance definition and expressions for a process in continuous and discrete time, along with the properties of its estimator.

Type	Autocovariance	
continuous	$c(\tau) := \mathrm{cov}\left[\underline{x}(t), \underline{x}(t+\tau)\right]$ where $\tau \in \mathbb{R}$ is the lag for a continuous time process (in time units)	(A14)
discrete	$c_{\mathrm{d}}^{(\Delta)}(j) := \dfrac{\Delta^2[j^2\gamma(j\Delta)]}{2\Delta[j^2]} = \dfrac{1}{2}\left((j+1)^2\gamma\left((j+1)\Delta\right) + (j-1)^2\gamma\left((j-1)\Delta\right) - 2j^2\gamma(j\Delta)\right)$ where $j \in \mathbb{Z}$ is the lag for the process at discrete time (dimensionless)	(A15)
classical estimator	$\hat{\underline{c}}_{\mathrm{d}}^{(\Delta)}(j) = \dfrac{1}{\zeta(j)}\sum_{i=1}^{n-j}\left(\underline{x}_i^{(\Delta,D)} - \dfrac{1}{n}\left(\sum_{l=1}^{n}\underline{x}_l^{(\Delta)}\right)\right)\left(\underline{x}_{i+j}^{(\Delta,D)} - \dfrac{1}{n}\left(\sum_{l=1}^{n}\underline{x}_l^{(\Delta)}\right)\right)$ where $\zeta(j)$ is usually taken as: n or $n-1$ or $n-j$	(A16)
expectation of classical estimator	$\mathrm{E}\left[\hat{\underline{c}}_{\mathrm{d}}^{(\Delta)}(j)\right] = \dfrac{1}{\zeta(j)}\left((n-j)c_{\mathrm{d}}^{(\Delta)}(j) + \dfrac{j^2}{n}\gamma(j\Delta) - j\gamma(n\Delta) - \dfrac{(n-j)^2}{n}\gamma\left((n-j)\Delta\right)\right)$	(A17)

Table A5: Variogram definition and expressions for a process in continuous and discrete time, along with the properties of its estimator.

Type	Variogram	
continuous	$v(\tau) := c(0) - c(\tau)$	(A18)
discrete	$v_{\mathrm{d}}^{(\Delta)}(j) := \gamma(\Delta) - c_{\mathrm{d}}^{(\Delta)}(j)$	(A19)
classical estimator	$\underline{\hat{v}}_{\mathrm{d}}^{(\Delta)}(j) = \underline{\hat{\gamma}}(\Delta) - \underline{\hat{c}}_{\mathrm{d}}^{(\Delta)}(j)$	(A20)
expectation of classical estimator	$\mathrm{E}\left[\underline{\hat{v}}_{\mathrm{d}}^{(\Delta)}(j)\right] = \mathrm{E}\left[\underline{\hat{\gamma}}(\Delta)\right] - \mathrm{E}\left[\underline{\hat{c}}_{\mathrm{d}}^{(\Delta)}(j)\right]$	(A21)

Table A6: Power spectrum definition and expressions for a process in continuous and discrete time, along with the properties of its estimator.

Type	Power spectrum	
continuous*	$s(w) := 4\int_0^\infty c(\tau)\cos(2\pi w\tau)\,\mathrm{d}\tau$	(A22)
discrete**	$s_{\mathrm{d}}^{(\Delta)}(\omega) := 2\Delta\gamma(\Delta) + 4\Delta\sum_{j=1}^{\infty} c_{\mathrm{d}}^{(\Delta)}(j)\cos(2\pi\omega j)$ where $\omega \in \mathbb{R}$ is the frequency for a discrete time process (dimensionless; $\omega = w\Delta$)	(A23)
classical estimator	$\underline{\hat{s}}_{\mathrm{d}}^{(\Delta)}(\omega) = 2\Delta\underline{\hat{c}}_{\mathrm{d}}^{(\Delta)}(0) + 4\Delta\sum_{j=1}^{n}\underline{\hat{c}}_{\mathrm{d}}^{(\Delta)}(j)\cos(2\pi\omega j)$	(A24)
expectation of classical estimator**	$\mathrm{E}\left[\underline{\hat{s}}_{\mathrm{d}}^{(\Delta)}(\omega)\right] = 2n\Delta\big(\gamma(\Delta) - \gamma(n\Delta)\big)/\zeta(0) +$ $+4\Delta\sum_{j=1}^{n}\frac{\cos(2\pi\omega j)}{\zeta(j)}\left((n-j)c_{\mathrm{d}}^{(\Delta)}(j) + \frac{j^2}{n}\gamma(j\Delta) - j\gamma(n\Delta) - \frac{(n-j)^2}{n}\gamma((\mathrm{n}-j)\Delta)\right)$	(A25)

*Equation (A22) can be solved in terms of c to yield (the inverse cosine Fourier transformation): $c(\tau) = \int_0^\infty s(w)\cos(2\pi w\tau)\,\mathrm{d}w$. Also, it can be solved in terms of γ to yield: $\gamma(m) = \int_0^\infty s(w)\frac{\sin^2(\pi w m)}{(\pi w m)^2}\mathrm{d}w$ and $s(w) = -2\int_0^\infty (2\pi w m)^2\gamma(m)\cos(2\pi w m)\,\mathrm{d}m$ (Koutsoyiannis, 2013a).

**Equations (A23) and (A25) are more easily calculated with fast Fourier transform (fft) algorithms. Also, Koutsoyiannis (2013a) shows how the discrete time power spectrum can be linked directly to the continuous time one, without the use of autocovariance function.