



Can a simple stochastic model generate rich patterns of rainfall events?

Simon-Michael Papalexiou^a, Demetris Koutsoyiannis^a, Alberto Montanari^{b,*}

^a Department of Water Resources, Faculty of Civil Engineering, National Technical University of Athens, Heron Polytechniou 5, GR-157 80 Zographou, Greece

^b Department DICAM, University of Bologna, Italy

ARTICLE INFO

Article history:

Received 22 December 2010

Received in revised form 14 June 2011

Accepted 10 October 2011

Available online 18 October 2011

This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Uwe Haberlandt, Associate Editor

Keywords:

Long-term persistence

Power-type tails

Rainfall modelling

SUMMARY

Several of the existing rainfall models involve diverse assumptions, a variety of uncertain parameters, complicated mechanistic structures, use of different model schemes for different time scales, and possibly classifications of rainfall patterns into different types. However, the parsimony of a model is recognized as an important desideratum as it improves its comprehensiveness, its applicability and possibly its predictive capacity. To investigate the question if a single and simple stochastic model can generate a plethora of temporal rainfall patterns, as well as to detect the major characteristics of such a model (if it exists), a data set with very fine timescale rainfall is used. This is the well-known data set of the University of Iowa comprising measurements of seven storm events at a temporal resolution of 5–10 s. Even though only seven such events have been observed, their diversity can help investigate these issues. An evident characteristic resulting from the stochastic analysis of the events is the scaling behaviors both in state and in time. Utilizing these behaviors, a stochastic model is constructed which can represent all rainfall events and all rich patterns, thus suggesting a positive reply to the above question. In addition, it seems that the most important characteristics of such a model are a power-type distribution tail and an asymptotic power-type autocorrelation function. Both power-type distribution tails and autocorrelation functions can be viewed as properties enhancing randomness and uncertainty, or entropy.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction and motivation

Rainfall has been traditionally regarded as a random process with several peculiarities, mostly related to intermittency and non-Gaussian behavior. However, many have been not satisfied with the idea of a pure probabilistic or stochastic description of rainfall and favored a deterministic modelling option. For example, Eagleson (1970, p. 184) states “The spacing and sizing of individual events in the sequence is probabilistic, while the internal structure of a given storm may be largely deterministic”. Such a perception of rainfall is also reflected in common engineering practices, such as the construction of design storms, in which the total depth may be determined by probabilistic considerations but the arrangement of rainfall depth increments follows a deterministic procedure, e.g. a pre-specified dimensionless hyetograph.

More recently, developments of nonlinear dynamical systems and chaos allowed many to apply algorithms from these disciplines in rainfall and claim for having discovered low dimensional deterministic dynamics in rainfall (see e.g., Sivakumar, 2000; Puente and Sivakumar, 2007). However, such results have been disputed by others (e.g., Schertzer et al., 2002; Koutsoyiannis, 2006a). In the latter study, among other data sets, a high temporal resolution data

record was used, in which the application of chaos detection algorithms did not give any indication of low dimensional chaos.

This high resolution record is one of seven storms that were measured by the Hydrometeorology Laboratory at the University of Iowa using devices that are capable of high sampling rates, once every 5 or 10 s (Georgakakos et al., 1994). This unique data set allows inspection of the rainfall process at very fine time scales and was the subject of several extensive analyses including multifractal analysis and multiplicative cascades (Cârsteanu and Foufoula-Georgiou, 1996) and wavelet analysis (Kumar and Foufoula-Georgiou, 1997). However, apart from such more technical analyses, this unique data set offers a basis for simpler yet more fundamental investigations that could provide insights for the characterization and mathematical modelling of the rainfall process; this will be attempted in the next sections. In this respect, the Iowa data set allows revisiting and acquiring better insight on the questions whether a single model can or cannot generate different types of events with enormous differences among them and, if yes, how such a model would look like. First, will it be deterministic or stochastic? A deterministic perception of the rainfall process may seem in accord to the high temporal dependence (autocorrelation) of the rainfall process at small lag times. However, this may indicate a misconception because at high autocorrelation without a specified underlying reason (an a priori known deterministic control) may increase rather than reduce uncertainty (Koutsoyiannis, 2010) and thus may require a

* Corresponding author.

E-mail address: alberto.montanari@unibo.it (A. Montanari).

stochastic description. In the latter case, fundamental behaviors to be explored are (a) the long (e.g., power-law) or short (e.g., exponential) tails in probability distribution function and (b) the long or short tails of the autocorrelation function. In both cases, long tails imply high uncertainty and may comply with the maximum entropy principle applied with certain constraints (Koutsoyiannis, 2005a,b).

It should be emphasized from the beginning that the scope of this paper is more explanatory than descriptive. In this respect, some general properties of a candidate rainfall modelling approach, rather than the construction of a complete and accurate model, are sought. Besides, as the empirical basis of this study is the Iowa data set which comprises only seven uninterrupted single storms, it is impossible to study all aspects of the rainfall process, and generalize the validity of our findings for other seasons or other locations. For example intermittency, a very important peculiarity of the rainfall process is left out of this study. For the latter, and especially its relationship to the maximum entropy principle, the interested reader is referred to a recent study by Koutsoyiannis (2006b).

2. General properties of rainfall data set

2.1. The data

Seven storm events of high temporal resolution, recorded by the Hydrometeorology Laboratory at the Iowa University (Georgakakos et al., 1994), are the data set of this study. The original measurements were taken every 5 or 10 s; however, for uniformity here we use the 10-s resolution for all events. Fig. 1 illustrates the patterns of the seven storms.

The events are characterized by a variable duration and also exhibit large statistical differences among them. Specifically, summary statistics like the mean, the standard deviation, the skewness and the kurtosis, shown in Table 1, differ notoriously among the events, up to two orders of magnitude (e.g., the kurtosis coefficient). In the following analyses, the different events are analyzed either separately or jointly. For the latter type of analysis, which is consistent with the scope of the paper to seek whether a single model can or cannot generate all different types, a merged sample of all events is used.

2.2. Scaling in state

The term scaling in state (Koutsoyiannis, 2005a) refers to the power-law behavior of the probability distribution of a process. Whether or not a natural process is characterized by a power-law distribution is of great importance, as a power-law process implies that extreme events are not only more frequent in comparison to an exponential-law process, but also more severe. Clearly, the frequency and the magnitude of extreme events in natural processes like rainfall, have many practical consequences, e.g., in the design of hydraulic works.

In practice, the identification and the characterization of a natural process as a power-law process is a difficult task. Natural processes that are considered to be power-law, do not exhibit a single power law distribution over the entire domain. Thus, the range over which the power-law holds, i.e. the distribution tail, must be identified and this is not trivial. Actually, inferences related to distribution tail that are based on sample data are uncertain. Therefore, in the best case, the validity of a power law might be conjectured, if the empirical data are consistent with the hypothesized power law and do not falsify the power-law hypothesis.

Generally, there are several methods for identifying power-law behavior in empirical data, e.g., methods based on least-square fitting or maximum likelihood, but none of them seems to be universally accepted (see e.g., Clauset et al., 2009 and references therein). Nevertheless, one of the most common procedures used for discerning power-law behavior in empirical data, which dates back to the end of 19th century in the works of Pareto, is based on least-square fitting.

Mathematically, a random variable X follows a power-law distribution, if its probability density function is of the form

$$f_X(x) \sim L(x)x^{-a-1} \quad (1)$$

where $a > 0$ is a constant known as the scaling exponent or the tail index, and $L(x)$ is a slowly varying function, that is a function satisfying $\lim_{x \rightarrow \infty} L(cx)/L(x) = 1$, where c is a constant. The essence of a slowly varying function is that asymptotically it does not affect the power-law behavior of the distribution, thus controlling the shape of the distribution only over a finite domain of values. Straightforwardly from (1), the q th moment of a power-law distribution, defined as $m_q := \int_{-\infty}^{\infty} x^q f_X(x) dx$, diverges if $q > a$.

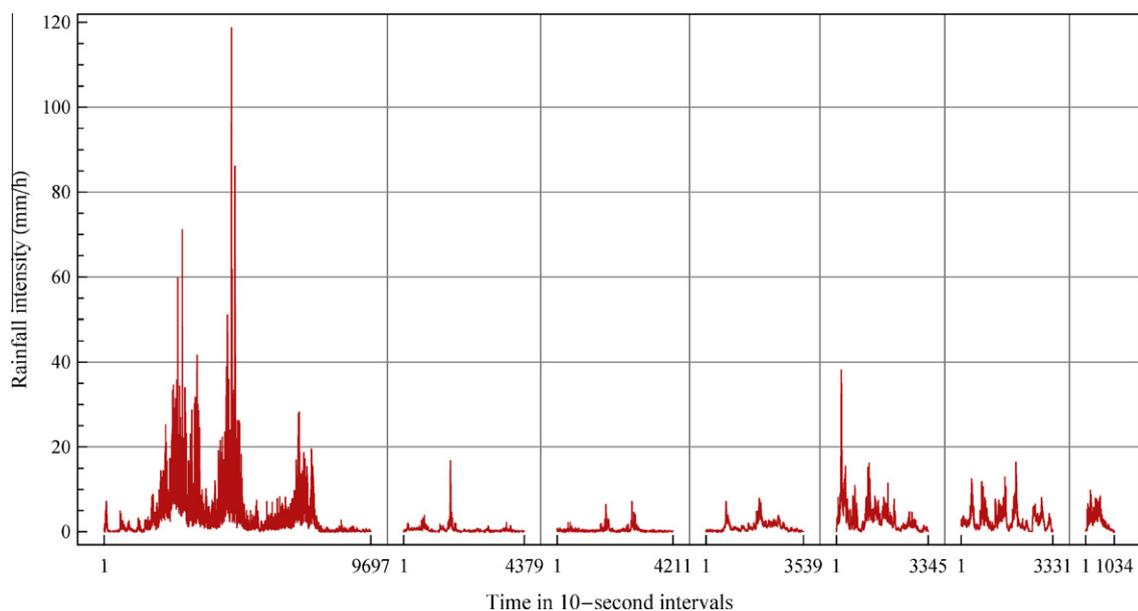


Fig. 1. The seven storm events recorded by the Hydrometeorology Laboratory at the Iowa University.

Table 1
Summary statistics of the seven storm events.

Event #	1	2	3	4	5	6	7	All
Sample size	9697	4379	4211	3539	3345	3331	1034	29,536
Mean (mm/h)	3.89	0.50	0.38	1.14	3.03	2.74	2.70	2.29
Standard deviation (mm/h)	6.16	0.97	0.55	1.19	3.39	2.20	2.00	4.11
Skewness	4.84	9.23	5.01	2.07	3.95	1.47	0.52	6.54
Kurtosis	47.12	110.24	37.38	5.52	27.34	2.91	-0.59	91
Hurst Exponent	0.94	0.79	0.89	0.94	0.89	0.87	0.97	0.89

It is also apparent from (1) that in a double-logarithmic plot, a power-law distribution (for both the probability density function and the probability distribution function) would be depicted as a straight line—at least in the range of values where the power law holds, i.e. the distribution tail. Thus, the slope of the least-square-fitted line to the tail of the empirical distribution (which, by virtue of (1) is proportional to x^{-a}) is an estimate of the state-scaling exponent. Using the aforementioned Pareto’s method, we depict in Fig. 2 the empirical probability distribution (constructed by using the Weibull plotting position) of the merged Iowa dataset and the least-square-fitted line to the empirical tail. A power law with $a \approx 3$ seems to describe the tail (at probability of exceedance smaller than 1%).

2.3. Scaling in time

Since Hurst (1951) empirically discovered scaling in time (Koutsoyiannis, 2005b), else known as long-term persistence (LTP), this same behavior has been identified in many other natural processes, as well as time series from many other scientific disciplines, e.g., in economy and in network traffic (e.g., Baillie, 1996; Leland et al., 2002). Ever since, LTP has been an active research field, as its importance necessitated not only theoretical accounts, but also, practical approaches concerning primarily the estimation of its strength and the development of models capable of generating synthetic time series with LTP behavior.

Basically, scaling in time can be defined in terms of the averaged process on several time scales k , i.e.

$$X^{(k)}(\tau) := \frac{1}{k} \sum_{t=(\tau-1)k+1}^{k\tau} X(t) \tag{2}$$

In a scaling process the following expression holds, i.e.,

$$[X^{(k)}(\tau) - \mu_X] \stackrel{d}{\sim} k^{H-1} [X(t) - \mu_X] \tag{3}$$

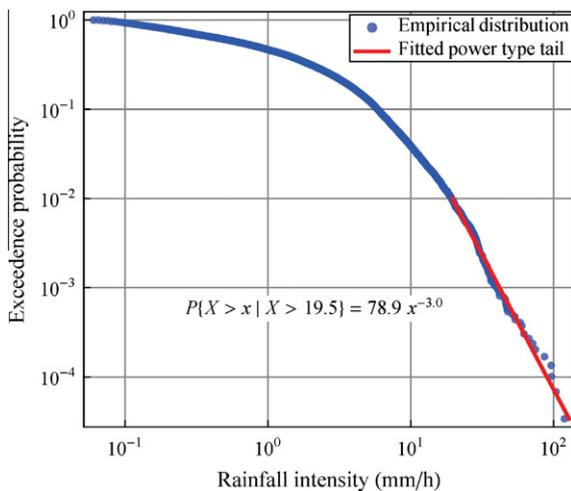


Fig. 2. Empirical probability distribution (Weibull plotting position) of the merged Iowa dataset and the least-square-fitted line to the empirical tail.

for any t and τ , where H is the scaling exponent or the so-called Hurst coefficient, and $\stackrel{d}{\sim}$ stands for equality in probability distribution. This process has recently been termed the Hurst-Kolmogorov process (HK; to give credit to Kolmogorov, 1940, who was the first to propose it). If X is Gaussian the process is also called fractional Gaussian noise (fGn), due to Mandelbrot and Van Ness (1968). As can be easily derived by (3), $\sigma_{X^{(k)}} = k^{H-1} \sigma_X$, that is, the aggregated process’s standard deviation is proportional to k^{H-1} and not to $k^{-0.5}$ as is in the case of independent processes. In addition, the autocorrelation function $\rho(\tau) \sim \tau^{2H-2}$ as $\tau \rightarrow \infty$ and the spectral density $S(\omega) \sim \omega^{1-2H}$. While in the HK process the property (3) holds for all time scales, in other processes it may hold only asymptotically, as scale tends to infinity. Again the Hurst coefficient H is an important characteristic of the asymptotic behavior. For example, in a Markovian process, $H = 0.5$ (as in independent processes).

With reference to LTP identification and parameter estimation—a non-trivial issue—many methods have been developed (e.g. based on maximum likelihood, the periodogram, the variance, the rescaled range and others concepts (e.g., Taquu et al., 1995; Taquu and Teverovsky, 1998; Tyralis and Koutsoyiannis, 2011), each having its advantages and drawbacks.

In this study, we chose to estimate the Hurst coefficients H of each of the seven storm events, and additionally of the merged dataset, by using a method that is based on the scaling property of the standard deviation, i.e., $\sigma_{X^{(k)}} = k^{H-1} \sigma_X$. Taking the logarithms, it follows that $\ln \sigma_{X^{(k)}} = (H - 1) \ln k + \ln \sigma_X$, and consequently, the aggregated sample standard deviation $\sigma_{X^{(k)}}$ versus the timescale k in a double-logarithmic plot, would be depicted as a straight line (at least in the timescale range where the scaling holds) and the estimated Hurst coefficient is $H = 1 + \eta$, where η is the slope of the fitted linear regression line.

The estimated Hurst coefficients of the seven storm events are presented in Table 1; the variation among the estimated coefficients is high, from 0.77 to 0.97 with a mean value 0.88. Nevertheless, under the assumption that the seven storm events can be considered as the realizations of a single process, a better estimate of the Hurst coefficient would result if the estimation is carried out on the merged dataset, taking care in the aggregation procedure that individual storm events do not interfere with each other. As Fig. 3 reveals, the scaling in the merged event seems to hold over the whole range of timescales, while the estimated Hurst coefficient is 0.94.

3. Stochastic analysis of the rainfall data set

3.1. The simulation scheme

As previously mentioned, the major target of this study is first to explore if the seven storm events could be considered as the outcome of a sole and simple stochastic process, and second, to identify the basic characteristics of this process. To this aim we proceeded heuristically; that is, we set a stochastic simulation scheme to generate synthetic rainfall series whose statistics are subsequently compared with those of the observed records. The aim is to check whether one cannot reject the hypothesis that the statistics themselves are coincident and therefore the observed

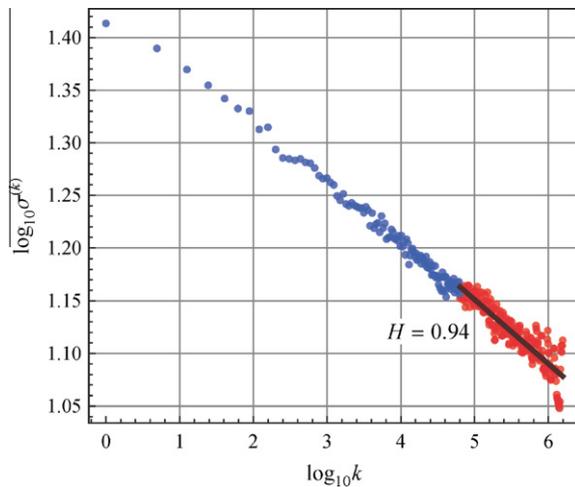


Fig. 3. Double logarithmic plot of sample standard deviation versus scale of averaging for the normalized merged event.

and synthetic records could be regarded as realizations of the same stochastic process. We considered different stochastic processes. In principle, candidate processes should include power law as well exponential solutions for both the marginal probability distribution and the autocorrelation. As mentioned above, there are two major questions that need to be answered: the first, concerns the scaling in state, i.e., whether or not, the stochastic process's marginal probability distribution is power type or exponential type. The second, concerns the scaling in time, i.e., whether or not, the autocorrelation structure is power type or exponential.

Regarding the marginal distribution, it is straightforward that realizations from a stochastic process with a power-type marginal distribution would exhibit large differences from an exponential marginal distribution, mainly because a power-type distribution assigns large probabilities to the extreme events, which signifies high variability and uncertainty. Clearly, this behavior is in agreement with the large variability observed in the seven recorded storm events, and in addition, the whole dataset does not falsify the power-law hypothesis of the marginal distribution (see Section 2.2). Consequently, the power-law hypothesis of the marginal distribution is accepted as rational and valid and only stochastic processes with power law marginal distribution were considered for the simulation.

In contrast to the choice of the marginal distribution, the a priori decision of a particular autocorrelation structure for the stochastic process is not simple. Short term persistence (STP) models have been a frequent choice in simulating natural phenomena, but they are often unjustifiably adopted (Koutsoyiannis and Montanari, 2007; Koutsoyiannis et al., 2009). In fact, an LTP autocorrelation structure, in many cases, may be more appropriate (see, for instance, Mandelbrot and Van Ness, 1968; Mandelbrot and Wallis, 1969). Additionally, it is not clear, how intensively the autocorrelation structure of a stochastic process—taking into account that the marginal distribution remains the same—affects the variability of the sample statistics among different realizations, e.g., the statistics among the simulated storm events addressed in this study. Thus, even in the case when the empirical evidence supports the adoption of a certain autocorrelation structure, and in view of the intrinsic uncertainty of this choice, it is valuable to perform a comparison of different scenarios, i.e., a comparison between STP and LTP autocorrelation structures. Therefore, this rationale suggests a side-by-side comparison between an STP model and an LTP model in view of the behaviors of the observed data.

The following sections present the simulation scheme which consists of the following seven steps: (1) application of an

appropriate normalizing transformation to the original dataset (Section 3.2); (2) analysis of the empirical ACF (Section 3.4); (3) identification and calibration of an STP model and an LTP model (Sections 3.5 and 3.6) to the normalized dataset; (4) correction of the model standard deviation bias (Section 3.7); (5) simulation of normal synthetic time series (Section 3.8); (6) generation of the synthetic rainfall time series by applying the inverse transformation (see Section 3.2) to the normal synthetic time series; and (7) statistical analysis of the synthetic time series (Section 4).

3.2. Normalizing the original data

The Gaussian or the Normal distribution is probably the most known and the most widely used distribution in statistics, with applications also in natural sciences. There are two theoretical reasons that justify the ubiquity of the Normal distribution in statistics and its application in other scientific fields. The first, relates to the central limit theorem (CLT) that—loosely speaking—states that the sum of independently and identically distributed (i.i.d) random variables tends to the Normal distribution as the number of summands tends to infinity. The second is the principle of maximum entropy (Jaynes, 1957), which states that, among all possible distributions with known mean and variance, the normal distribution is the one that maximizes the Boltzmann–Gibbs–Shannon information entropy (see also Shannon Claude and Weaver, 1948).

Nevertheless, it seems that geophysical data are seldom normal. Empirical data show that many geophysical processes, like rainfall and river discharge, may depart mildly or severely from normality, especially at small time scales. A relevant example is the dataset addressed in this study. Specifically, departures from normality may be identified in skewness, e.g., positively or negatively skewed empirical data, in the asymptotic behavior of the distribution tail, e.g., a stretched exponential tail or a power-type tail, and of course, in the variable's domain. Thus, as there exist theoretical reasons that favor normality in many cases, theoretical reasons also exist that do not support it (see e.g., Koutsoyiannis, 2005a).

For instance, it is well known that a normal variable ranges over the whole real axis, while many natural processes are positively defined, that is, have a lower limit at zero, while a solid reason to fix an upper limit very rarely exists. While the previous reasons explain why departures from normality are so common in nature, a formal and generalized method for simulating non-normal data with a certain autocorrelation structure does not exist, although heuristic solutions were frequently proposed (for a hydrological example, see Montanari et al., 1997). In contrast, several methods exist addressing the simulation of normal data with STP or LTP autocorrelation structures (e.g., Box et al., 1994; Koutsoyiannis, 2000; Brockwell and Davis, 2009). A common technique for simulating non-normal data consists of transforming the non-normal dataset to normal, by applying a normalizing transformation, next, simulating normal data by implementing a standard model, and finally, de-normalizing the normal data by applying the inverse transformation. Basically, this is the methodology followed also in this study, which presents the inconvenience that finding an appropriate normalizing transformation is not always a trivial task, and clearly, a general method for normalizing all types of data does not exist. It is well known that there are some general and commonly used families of transformations, like the Box-Cox family of transformations (Box and Cox, 1964), that in many cases give satisfactory results. Unfortunately, such general and simple transformations were not effective for the case of the Iowa dataset. In particular, while the application of the Box-Cox transformation resulted in approximately normal data for the upper empirical tail, it failed to normalize the lower tail, namely the values near zero. A frequently used solution to solve this problem is the normal quantile transform (also called normal quantile score; Kelly and

Krzysztofowicz, 1997) which, however, is an empirical transformation that is defined over the range of the observed data only and cannot be extrapolated.

Therefore, in order to normalize the Iowa dataset we introduce here heuristically (by extending a transformation by Koutsoyiannis et al., 2008) a five-parameter normalizing transformation given by

$$z(t) = g(x(t)) = [\alpha x(t)^{-\zeta} + \beta] \left(\gamma + \sqrt{\left(1 + \frac{1}{\delta}\right) \ln\{1 + \delta[x(t) - \gamma]^2\}} \right) \quad (4)$$

where $z(t)$ and $x(t)$ are the transformed and original values of the rainfall intensity, which are realisations of the stochastic processes $Z(t)$ and $X(t)$, respectively, and $\alpha, \beta, \gamma, \delta, \zeta$ are the parameters to be estimated. The two factors of the product in the right hand side are introduced to normalize the lower and the larger values, respectively.

While this transformation was identified heuristically, its construction was based on two theoretical aspects. First, equation (4) ensures that the random variable $Z \sim N(0, 1)$ ranges from $-\infty$ to ∞ . Obviously, inspection of (4) reveals that for $\{\alpha, \beta, \delta, \zeta\} \in (0, \infty)$ and $\gamma \in (-\infty, 0)$, the random variable $Z \in (-\infty, \infty)$, as $\lim_{x \rightarrow 0^+} g(x) = -\infty$ and $\lim_{x \rightarrow \infty} g(x) = \infty$. Second, the probability density function (pdf) of the random variable X should be long tailed as the empirical evidence supports this assumption (see Section 2.2). Again, inspection of (4) reveals that for large values of x , $g(x) \sim [2\beta^2 (1 + 1/\delta) \ln x]^{0.5}$, and taking into account that $f_Z(z) \sim \exp(-z^2/2)$ and combining the two equations, we get $f_X(x) \sim f_Z(g(x)) \sim x^{-\beta^2(1+1/\delta)}$ and thus the pdf of the variable X is long tailed.

Finally, we estimated the parameters of (4) for the transformed merged Iowa dataset by using the method of least-squares, and particularly, we numerically minimized the sum of squared errors between values of the standardized normal variate that correspond to the values of the empirical normal distribution (obtained by applying the normal quantile transformation) and the respective values result from the application of (4) to the original rainfall values. The resulted estimates were $\alpha = 0.41$, $\beta = 2.49$, $\gamma = -2.13$, $\delta = 4.09$ and $\zeta = 1.18$. The transformed data in comparison with the original data are presented in Fig. 4. Clearly, as the Fig. 4 demonstrates, the transformed data are satisfactorily normalized.

3.3. Identification and calibration of the stochastic models

A Gaussian (normal) stochastic process is completely characterized when its second-order distribution, i.e., $F_X(x_i, x_j; t_i, t_j) = P\{X(t_i) \leq x_i, X(t_j) \leq x_j\}$ for any $i \neq j$, is known. Normalizing the marginal distribution of a stochastic process by a transformation, does not necessarily result in jointly normal distribution (Feller, 1971, p.70). However, it is important to check if a particular, marginally normalized, data has also become Gaussian in terms of the multivariate joint distribution or not. A rough indication of joint normality is provided by the linear relation of conditional expectation of a variable X_i given X_j for $i \neq j$. Fig. 5 depicts the normalized rainfall intensity versus the 1-time-step and 10-time-step shifted normalized rainfall intensity. It can be seen that the empirical points are spread around a straight line, which is an indication of joint normality. This linearity should not be regarded as a surprise, given that it is consistent with the principle of maximum entropy applied on a multivariate setting with constraints of known mean, variance and lag-1 autocorrelation.

As discussed in Section 2.3, scaling in time exists and is quantified by an estimated Hurst coefficient $H = 0.94$. Similarly, analysis of the transformed dataset, using the same methods as in Section 2.3, reveals also a high value of the Hurst coefficient, i.e., $H = 0.92$. Thus,

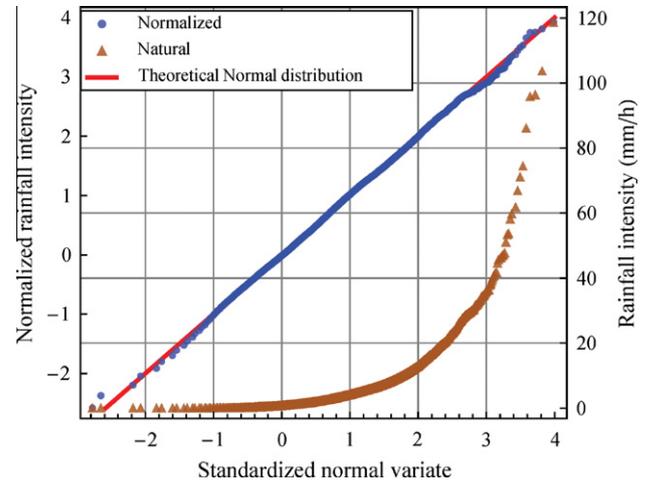


Fig. 4. Probability plot of the natural (recorded) and the normalized rainfall intensity data.

if we accept the assumption of scaling in time, a serious issue arises; that is, almost all classical estimators of statistics (exception is the mean value) are highly biased (e.g., Koutsoyiannis, 2003). So in order to set up an accurate and consistent stochastic model to simulate a normal process—that is, a model that sufficiently reproduces the mean, the standard deviation and the autocorrelation structure of the observed sample—unbiased and accurate estimates of the aforementioned statistics are necessary.

3.4. Empirical autocorrelation function (ACF)

It is well known, that for finite samples the typical estimate $\hat{\rho}_l$ of the lag- l autocorrelation is a biased estimator of the true autocorrelation ρ_l and the more intense the autocorrelation structure is the more biased the estimator becomes. In particular, in the presence of scaling in time the bias can be corrected by the following formula (see Koutsoyiannis, 2003 and the references therein),

$$\tilde{\rho}_l = \hat{\rho}_l \left(1 - \frac{1}{n^{2-2H}} \right) + \frac{1}{n^{2-2H}} \quad (5)$$

where $\tilde{\rho}_l$ stands for the unbiased estimator and H is the Hurst coefficient.

In this study, we used the unbiased estimator in (5) to estimate the empirical autocorrelation coefficients. We clarify, that to estimate $\hat{\rho}_l$ and consequently to estimate the unbiased estimator given in (5), we used the transformed merged sample that comprises the seven transformed storm events. This is a reasonable choice if we reckon the seven events as the outcome of a single process; and thus, while the empirical ACF may differ among events, all events share the same theoretical ACF. Furthermore, we note that we took special care in the estimation of the covariance in order to avoid overlapping among the events; specifically, we eliminated all products of the form $(x_t - \hat{\mu}_X)(x_{t-1} - \hat{\mu}_X)$ when x_t and x_{t-1} do not belong in the same storm event, and adjusted accordingly the number n of the sample size. The estimated unbiased empirical ACF—given a Hurst coefficient equal to $H = 0.92$, and for lags approximately up to 1000—is depicted in Fig. 6. Clearly, as Fig. 6 attests, the empirical autocorrelation structure is very intense, and particularly, the values of the small-lag autocorrelation coefficients are near to 1, while for lags near to 1000 the values are as high as 0.85.

3.5. The short-term persistence model

Probably, the most common STP stochastic model is the lag-one autoregressive model AR(1). This model belongs to the general

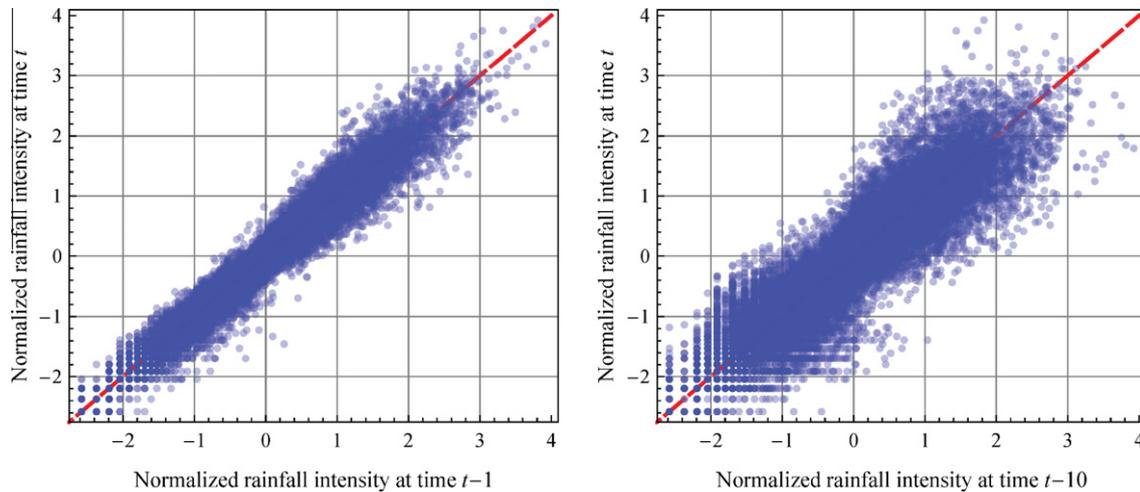


Fig. 5. Scatter plot of normalized rainfall intensity for time lags 1 and 10.

family of stochastic models known as autoregressive moving-average models ARMA(p,q)—comprehensively presented in Box et al. (1994). It is important to note that the ARMA(p,q) family, and especially the AR(1) model are not able to reproduce the scaling behavior in time or to preserve the Hurst coefficient (e.g., Koutsoyiannis, 2002). Consequently, they may be inappropriate for simulating natural phenomena exhibiting LTP.

Nevertheless, while from a theoretical viewpoint ARMA(p,q) models are considered STP models, for increasing values of the autoregressive and moving average order p and q they can provide very good approximations of the LTP structure and thus manage to reproduce, from a practical point of view, the scaling in time or to preserve the Hurst coefficient at least for small sample sizes (Papalexiou, 2007). It is clear, though, that high order ARMA(p,q) models are not parsimonious, i.e., many parameters need to be estimated therefore increasing the estimation variance.

Here, we chose the ARMA(2,2) model for the simulation of the normalized rainfall intensity. It is a model frequently used in hydrology that is able to generate time series that preserve the mean value μ_x , the variance σ_x^2 and the first four autocorrelation coefficients $\{\rho_1, \rho_2, \rho_3, \rho_4\}$. The stochastic process $\{X(t), t \in T\}$ that results from an ARMA(2,2) model is defined by

$$X(t) = a_1X(t - 1) + a_2X(t - 2) + \beta_1\varepsilon(t - 1) + \beta_2\varepsilon(t - 2) + \varepsilon(t) \quad (6)$$

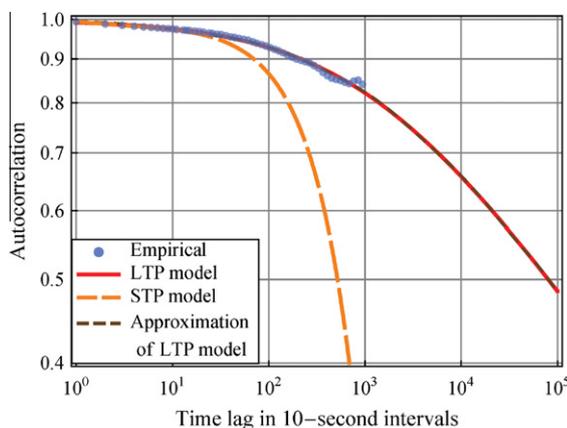


Fig. 6. Empirical ACF of the normalized merged event (corrected for bias), theoretical ACF of the fitted STP model, fitted power-type ACF given in eq. (9), and approximation of the latter by the sum of five AR(1) processes.

where $\{a_1, a_2, \beta_1, \beta_2\}$ are parameters, and $\varepsilon(t)$ is a normal white-noise process, i.e. consisting of independently, identically and normally-distributed random variables with mean $\mu_\varepsilon = 0$ and variance σ_ε^2 . Using typical estimation methods (Box et al., 1994), the resulting parameters for the transformed merged Iowa dataset are $\{a_1 = 1.51, a_2 = 0.51, \beta_1 = -0.57, \beta_2 = -0.19, \sigma_\varepsilon = 0.11\}$.

Once the model parameters are estimated, the theoretical ACF of the ARMA(2,2) for lags $\tau \geq 3$ degenerates to the ACF of an AR(2), i.e., $\rho_\tau = \alpha_1\rho_{\tau-1} + \alpha_2\rho_{\tau-2}$ and thus can be calculated recursively. Fig. 6 depicts the theoretical ACF of the fitted ARMA(2,2) model in comparison with the empirical ACF. Clearly, it preserves the first four autocorrelation coefficients, as expected, and also performs well for lags up to 50. Nevertheless, for higher lags, it clearly deviates from the empirical ACF as the exponential character of the theoretical ACF unfolds.

3.6. The long-term persistence model

Since the time when Hurst (1951) discovered the LTP behavior, the necessity to consistently simulate natural phenomena that exhibit LPT has led to the development of several stochastic processes and algorithmic procedures that reproduce the LTP behavior. Among the most common models are several algorithmic approximations of the HK (or fGn) process by Mandelbrot and Wallis (1969) Mandelbrot (1971), O’Connell (1974), Koutsoyiannis (2002), and the FARIMA(p,d,q) models introduced by Granger and Joyeux (1980) and Hosking (1981), that have gained popularity mainly in the last decade (for an application to hydrology see Montanari et al., 1997).

The theoretical ACFs of the HK and FARIMA(0, d ,0) processes are

$$\rho_\tau^{\text{FGN}} = \frac{1}{2} (|\tau - 1|^{2H} - 2|\tau|^{2H} + |\tau + 1|^{2H}) \sim \tau^{2H-2} \quad (7)$$

$$\rho_\tau^{\text{FARIMA}} = \frac{\Gamma(1-d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(\tau+1-d)} \sim \tau^{2d-1} \quad (8)$$

respectively. Clearly, the ACFs of those two models are asymptotically coincident, with $d = H - 1/2$, as (7) and (8) attest, whereas, time series generated by both of them preserve the scaling exponent H . Moreover, while the HK process model is a very simple model—essentially is one-parameter model, the FARIMA(p,d,q) models are much more flexible as the orders of p and q controls the STP behavior of the model.

Here we used a simple yet general approach to simulate LTP, obtained by approximating the real process with the sum of five independent AR(1) processes (note that Koutsoyiannis (2002) has shown that good approximations can be obtained even with summing three independent AR(1) processes). The implementation comprises two steps: first; we fit a generalized power-type (GP) ACF to the empirical ACF (see 3.4) and second; we approximate the fitted ACF by the ACF obtained as the sum of five independent AR(1).

Regarding the first step of our approach, we fit a theoretical ACF (consistent with the empirical evidence) to the empirical ACF in order to be able to extrapolate the correlation coefficients for lags as high as desired, instead of being confined in the lag-range provided by the estimated empirical ACF.

Here, we use a theoretical three-parameter power-type ACF that has the form (similar to Gneiting and Schlather, 2004)

$$\rho_{\tau}^{GP} := [1 + c(\frac{\tau}{a})^b]^{-\frac{1}{c}} \quad (9)$$

where $\{a > 0, b > 0, c > 0\}$ are parameters. The form of (9) can be considered as a natural generalization of an exponential ACF as the $\lim_{c \rightarrow 0} \rho_{\tau}^{GP} = \exp(-\tau/a)$. Asymptotically (9) behaves as $\rho_{\tau}^{GP} \sim t^{-b/c}$ and therefore, (9) and (7) possesses the same asymptotic behavior if $b/c = 2(1 - H)$. As a result, the fitted ρ_{τ}^{GP} would be consistent with the estimated $H = 0.92$ if $b/c = 0.16$. Thus, we fit the ρ_{τ}^{GP} by minimizing the square error between the ρ_{τ}^{GP} and the empirical ACF and by setting as a constrain $b/c = 0.16$. The estimated parameters are $\{a = 12881, b = 0.51, c = 3.18\}$. The fitted ρ_{τ}^{GP} is depicted in Fig. 6, which shows that the fit is satisfactory.

Turning to the second step of the LTP simulation procedure mentioned above, we decided to use an LTP model made up by the sum of five independent AR(1) process by following the idea that was first introduced by Mandelbrot (1971), to approximate the HK process. The same method was used by Koutsoyiannis (2002), for the same purposes, while Mudelsee (2007) proved empirically that the sum of n inflows generated by an AR(1) model in a river network, with n sufficiently large, ends up with a collective river discharge that exhibits LTP behavior.

Therefore the LTP model that was used herein to simulate the normalized rainfall intensity is given by

$$Y(t) = \sum_{i=1}^5 Y_i(t) \quad (10)$$

where $Y_i(t) = a_i Y_i(t - 1) + \varepsilon_i(t)$ is the i -th AR(1) process with mean $\mu_{Y_i} = 0$, variance σ_{Y_i} , lag-one autocorrelation coefficient a_i and $\varepsilon_i(t)$ is a normal white-noise process, with mean $\mu_{\varepsilon_i} = (1 - a_i)\mu_{Y_i} = 0$ and variance $\sigma_{\varepsilon_i}^2 = (1 - a_i^2)\sigma_{Y_i}^2$. Under the assumption of independence of the five AR(1) processes it can be easily proven that the theoretical ACF of (10) is given by

$$\rho_{\tau}^{LTP} = \sum_{i=1}^5 a_i^{\tau} \sigma_{Y_i}^2, \quad \text{with} \quad \sum_{i=1}^5 \sigma_{Y_i}^2 = 1 \quad (11)$$

The parameters of the five independent AR(1) processes were estimated by minimizing the square error between Eqs. (9) and (11) for lags as high as 10^4 . The resulting estimates are $\{a_1 = 0.9943, \sigma_{Y_1}^2 = 0.075\}$, $\{a_2 = 0.8719, \sigma_{Y_2}^2 = 0.029\}$, $\{a_3 = 0.999932, \sigma_{Y_3}^2 = 0.179\}$, $\{a_4 = 0.9994, \sigma_{Y_4}^2 = 0.138\}$ and $\{a_5 = 0.999998, \sigma_{Y_5}^2 = 0.578\}$. As the Fig. 6 reveals, the fitted ρ_{τ}^{LTP} , up to the lag- 10^4 , is satisfactory.

3.7. The standard deviation bias

One issue in stochastic modelling that may have serious consequences on the validity and accuracy of the simulation, and is often

neglected, concerns the differences in statistics that may occur between the theoretical process and its realizations. While the estimate of the mean is unbiased regardless of the dependence structure, this does not hold for the standard deviation. In fact, it is well known that the standard estimator S of the standard deviation is slightly biased even in the case of normally distributed and independent data (e.g., Bolch, 1968). However, the bias may become very large in a time dependent process as it increases monotonically with the increase of the autocorrelation. For certain known ACFs, like the one of the HK process, unbiased estimators have been developed (see Koutsoyiannis, 2003 and references therein).

In order to assess the standard deviation bias in random samples generated by the STP and the LTP models described in Sections 3.5 and 3.6, and for several different sample sizes, we performed a Monte Carlo simulation. Specifically, at first, 5000 independent samples were generated by each model and for several sample sizes, and in turn, a standard deviation correction factor was calculated, defined by $c_{SD} := \sigma/E[S]$ where σ is the true standard deviation of the STP and the LTP models, chosen as 1 in our simulations and $E[S]$ is calculated by Monte Carlo simulation.

The results, that are depicted in Fig. 7, are remarkable, especially in the case of the LTP model. In fact, the bias correction factors, for a sample size of 1000, are as high as 1.9 and 3.1 for the STP and the LTP models, respectively, while even for a very large sample size equal to 50,000, in the LTP case, the correction factor sustains a value of 1.7. Given that the correction factor depends of the sample size, the choice of the appropriate correction factor should be carried out by considering the number of the data generated with the simulation. Given that the normalizing transformation was applied to a sample of 29,536 values that comprised the seven storm events, it follows that we imposed a unit standard deviation to that complete sample. Consequently, all samples generated in this study, irrespective of their size, were multiplied by the correction factor that corresponds to a size of 29,536, that is, $c_{SD} = 1.04$ for the STP model and $c_{SD} = 1.81$ for the LTP model. In this way the correction to the standard deviation was imposed depending on the sample size that was used to constrain the standard deviation itself during the normalizing transformation.

3.8. Sample size and number of samples

As shown in Table 1, the seven recorded storm events have all different sample lengths varying from 1034 to 9697 values. In

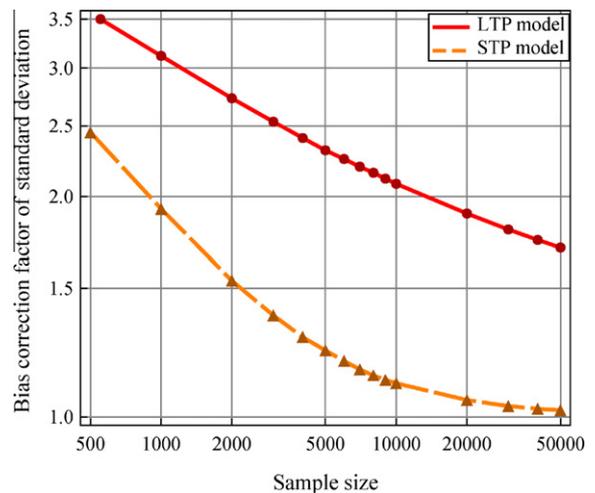


Fig. 7. Standard deviation bias correction factors for the STP and the LTP models for various sample sizes (dots and triangles) calculated by Monte Carlo simulation.

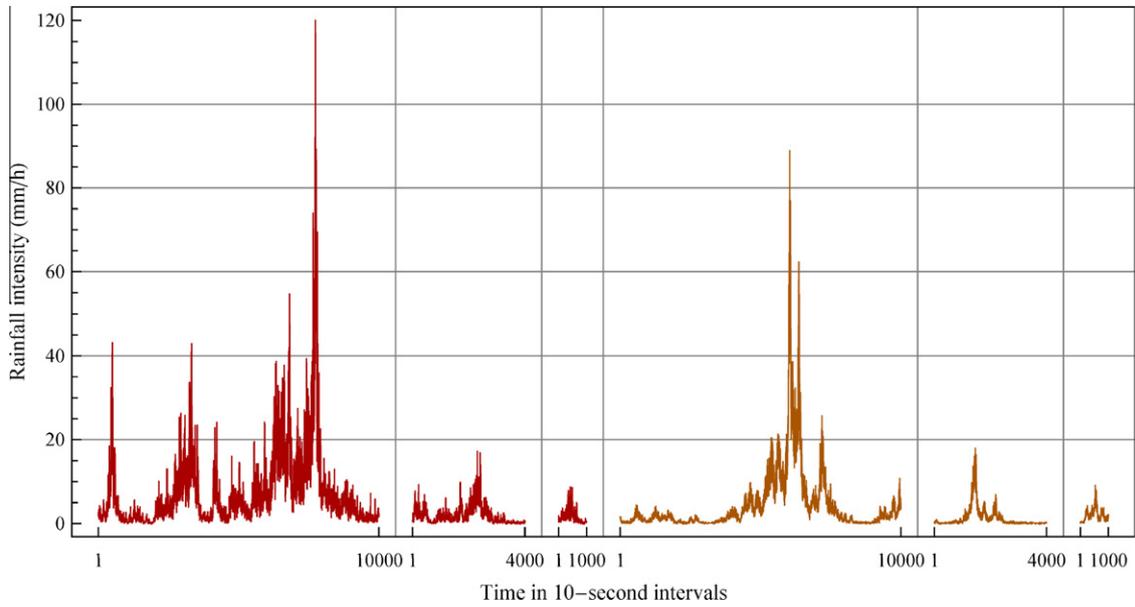


Fig. 8. Synthetic rainfall events generated by the LTP (the first three) and the STP (the last three) models for three characteristic sample sizes.

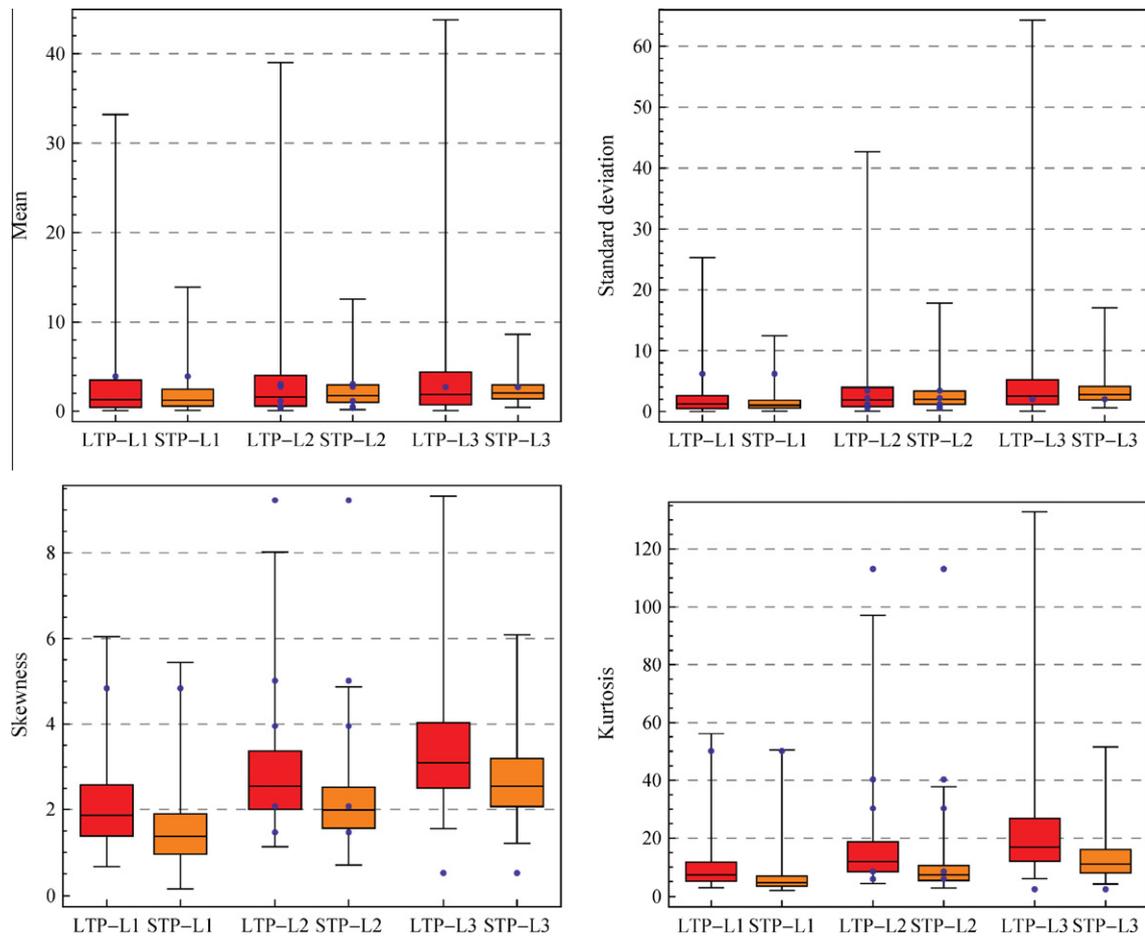


Fig. 9. Box plots of sample statistics estimated from the synthetic rainfall events generated by the LTP and STP models for the three characteristic sample sizes L1, L2 and L3. The dots represent the empirical points of the seven rainfall events.

order to compare the observed statistics with those as the synthetic series, it would be appropriate that the simulations have the same length of the observed records. For practicality only 3

sample sizes were used namely: 1000 (L1), which is very close to the size of event 7; 4000 (L2), close to the size of events from 2 to 6; and 10,000 (L3) representing event 1.

We generated 10,000 synthetic series for each sample length and for each model. Finally, we calculated for every synthetic series the mean, the standard deviation, the skewness, the kurtosis and the autocorrelations, which were compared with the respective statistics of the observed records.

4. Results of the stochastic simulation

Fig. 8 reports an example of visualized simulated events for the three different sample sizes (L1, L2 and L3) considered here and the two different models (three events generated by the LTP model on the left and three by the STP model on the right). Some differences in the patterns generated by the models are visible. For example, the pattern of the LTP model is characterized by a higher variability (although the marginal distributions are the same). By comparing the patterns with those of the observed records, which are shown in Fig. 1, one may notice that the variability of the observed record looks better reproduced by the LTP model.

Fig. 9 shows box plots of selected statistics computed on the simulated data (in this case also by referring to the original probability distribution), namely, mean, standard deviation, skewness and kurtosis. The observed statistics are also shown with dots. The box and the whiskers encompass 50% and 99%, respectively, of the computed statistics, while the median is indicated by a horizontal straight line. The box plots clearly show the different behaviors of the two models. Looking at the mean value, one should note that, not surprisingly, the two models are characterized by nearly the same median of the mean, but the variability in the LTP model is higher. Also expected is the higher variability of the standard deviation, skewness and kurtosis that is depicted in the other box plots. One may note that the LTP model is more skewed than the STP one. This result is explained by the higher

variability of a process (rainfall) that is bounded at zero. In general one may note that the higher uncertainty of the LTP model makes the fit more satisfactory, even though the observed points are very few and therefore do not allow more than a qualitative assessment.

Fig. 10 shows a comparison between the observed autocorrelation functions with those simulated by the models. First of all, one notes that the autocorrelation coefficients of the LTP model are higher in the tail of the autocorrelation function. This result is expected. Another relevant feature is the higher correlations showed by the STP model for low lags. This result, which is not intuitive, is due to the fact that the STP model, in order to reach a better fit of the tail of the ACF, reacts by increasing also the autocorrelation coefficients for low lags. Conversely, the power law behavior of the LTP model allows one to reach a better fit of the tail of the ACF without increasing much the correlation for low lags. Even in this case, the autocorrelation function of the LTP model appears to be more convincing in view of the observed pattern. One should note that this assessment is again qualitative in view of the small number of observed events.

However, apart from the comparison between the two models, one relevant conclusion is that both models look able to provide, within a relatively simple framework, a satisfactory fit of the observed behaviors.

5. Conclusions and discussion

Summarizing the above investigations, it can be said that a single and rather simple stochastic model can represent all rainfall events and all rich patterns appearing in each of the separate events making them look very different from one another. From a practical view point, such a model is characterized by high autocorrelation at fine scales, slowly decreasing with lag, as well as by

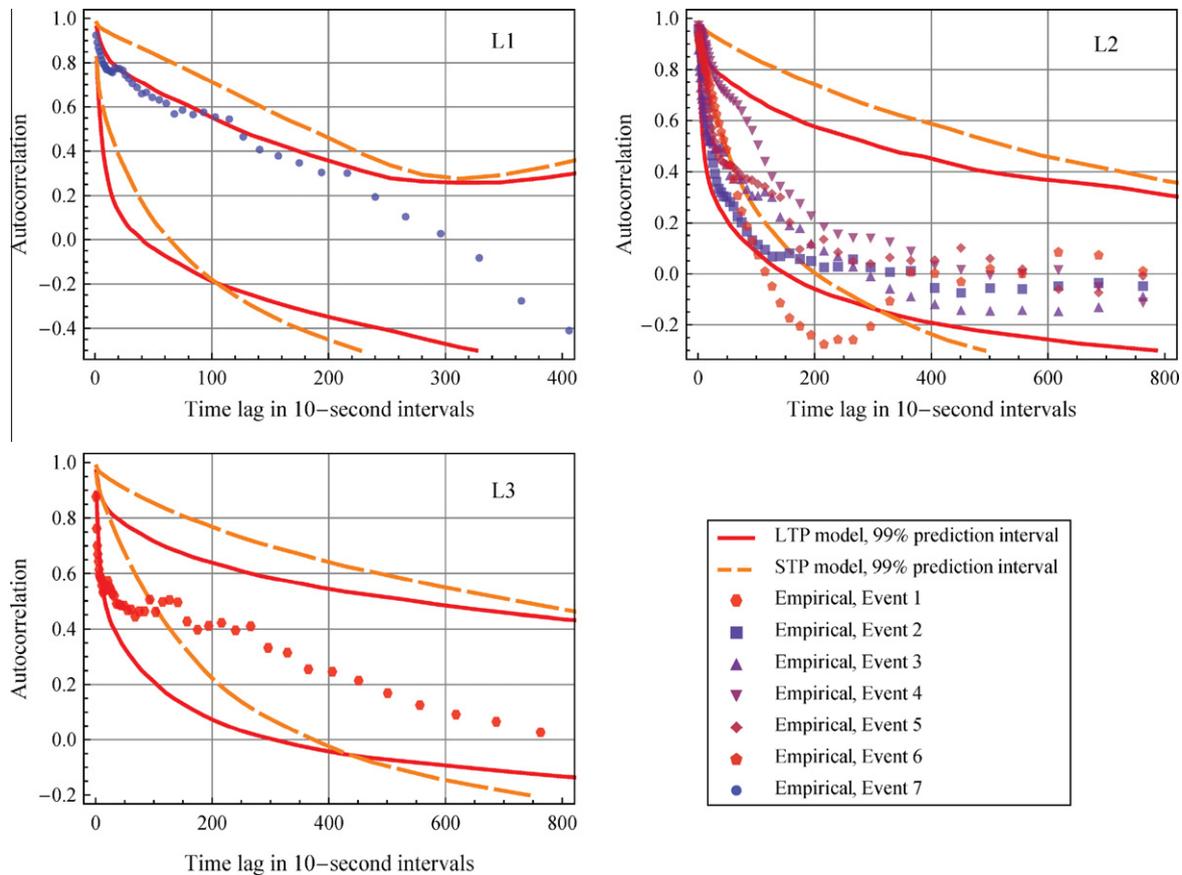


Fig. 10. Empirical autocorrelation functions of the seven rainfall events and 99% prediction intervals of ACF for the LTP and STP models.

distribution tails slowly decreasing with rainfall intensity. Such an autocorrelation form can indeed produce huge differences among different events and such a distributional form can produce enormously high rainfall intensities at times. Both these behaviors are just opposite to the more familiar processes resembling Gaussian white noise, which would produce very “stable” events with infrequent high intensities. In this respect, both high autocorrelations and distribution tails can be viewed as properties enhancing randomness and uncertainty (or entropy).

Whether the tails of both the marginal distribution and autocorrelation functions are long (meaning that are described by power-law functions) is difficult to conclude based merely on the data set of this study. Both these power-law functions are by definition asymptotic properties, and the exponents of power laws are theoretically defined for state or lag tending to infinity. In this respect, it seems impossible to verify such asymptotic laws by empirical studies, which necessarily imply finite sample sizes. But it is important that the empirical evidence presented in the current study does not falsify the hypothesis that both tails are long. Other empirical studies published recently do not falsify this hypothesis as well. Thus, Koutsoyiannis (2004a, 2004b) used daily rainfall records from numerous gauges worldwide, each covering a period longer than 100 years, and showed that the observed behaviors are consistent with the long distributional tail hypothesis.

If the hypothesis of a long tail of the distribution function is accepted, it seems that this can be quantified by an exponent α of about 3, which implies that only the first three moments of the distribution exist whereas all others are infinite. If the hypothesis of a long tail of the autocorrelation function is accepted, it seems that this can be quantified by a Hurst coefficient H as high as 0.94. Based on these findings, the construction of a stochastic model admitting asymptotically long tails from the outset seems a reasonable choice. After all, in a dynamical systems context, even the randomness is an asymptotic property per se, in the sense that it implies an infinite number of degrees of freedom. The fact that an infinite number of degrees of freedom cannot be verified (and perhaps neither falsified) empirically, does not preclude us from successfully using probabilistic descriptions and stochastic models of several processes including rainfall.

As mentioned earlier, long tails can be viewed as an enhancement of randomness and uncertainty in these processes. In the framework of this enhanced randomness, it seems to be useless to analyze each rainfall event separately as an attempt to infer dynamics of rainfall. Such an attempt, even using sophisticated methods such as wavelets, can perhaps be paralleled with one's attempt to explain the dynamics of the tossing of a coin by observing a series of “heads” and “tails”. In both cases, it may be misleading to seek substantial information in extremely random occurrences. A more useful target for such cases would be to elevate from the obscurity the underlying randomness and seek its own laws – and this was the main target of this paper.

Acknowledgment

We thank Geoff Pegram and one additional anonymous reviewer for their encouraging and constructive comments on two subsequent versions of the paper. We thank as well one anonymous reviewer for providing negative comments on the first version.

References

- Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.
- Bolch, B.W., 1968. More on unbiased estimation of the standard deviation. *American Statistician* 22, 27.
- Box, G.E., Cox, D.R., 1964. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)* 211–252.
- Box, G.E., Jenkins, G.M., Reinsel, G.C., 1994. *Time Series Analysis: Forecasting and Control*, 3rd ed. Prentice Hall, New Jersey.
- Brockwell, P.J., Davis, R.A., 2009. *Time Series: Theory and Methods*. Springer Verlag.
- Cărsteanu, A., Foufoula-Georgiou, E., 1996. Assessing dependence among weights in a multiplicative cascade model of temporal rainfall. *Journal of Geophysical Research* 101, 26363.
- Clauset, A., Shalizi, C.R., Newman, M.E., 2009. Power-law distributions in empirical data. *SIAM Review* 51, 661–703.
- Eagleson, P.S., 1970. *Dynamic Hydrology*. McGraw-Hill Inc., New York.
- Feller, W., 1971. *An introduction to probability theory and its applications*, 2nd ed. John Wiley & Sons Inc, New York.
- Georgakakos, K.P., Carsteanu, A.A., Sturdevant, P.L., Cramer, J.A., 1994. Observation and analysis of Midwestern rain rates. *Journal of Applied Meteorology* 33, 1433–1444.
- Gneiting, T., Schlather, M., 2004. Stochastic models that separate fractal dimension and the Hurst effect. *Society for Industrial and Applied Mathematics Review* 46 (2), 269–282.
- Granger, C.W.J., Joyeux, R., 1980. An introduction to long-memory time series and fractional differencing. *Journal of Time Series Analysis* 1, 15–29.
- Hosking, J.R., 1981. Fractional differencing. *Biometrika* 68, 165–176.
- Hurst, H.E., 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* 116, 770–808.
- Jaynes, E.T., 1957. Information theory and statistical mechanics. II. *Physical Review* 108, 171–190.
- Kelly, K.S., Krzysztofowicz, R., 1997. A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydraulics* 11, 17–31.
- Kolmogorov, A.N., 1940. Wiener'sche Spiralen und Einige Andere Interessante Kurven in Hilbertschen Raum. *Doklady Akademii nauk URSS* 26, 115–118.
- Koutsoyiannis, D., 2002. The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrological Sciences Journal* 47, 573–596.
- Koutsoyiannis, D., 2003. Climate change, the Hurst phenomenon, and hydrological statistics. *Hydrological Sciences Journal* 48, 3–24.
- Koutsoyiannis, D., 2004a. Statistics of extremes and estimation of extreme rainfall, 1, Theoretical investigation. *Hydrological Sciences Journal* 49, 575–590.
- Koutsoyiannis, D., 2004b. Statistics of extremes and estimation of extreme rainfall, 2, Empirical investigation of long rainfall records. *Hydrological Sciences Journal* 49, 591–610.
- Koutsoyiannis, D., 2005a. Uncertainty, entropy, scaling and hydrological stochasticity. 1. Marginal distributional properties of hydrological processes and state scaling. *Hydrological Sciences Journal* 50, 381–404.
- Koutsoyiannis, D., 2005b. Uncertainty, entropy, scaling and hydrological stochasticity. 2. Time dependence of hydrological processes and time scaling. *Hydrological Sciences Journal* 50, 1–426.
- Koutsoyiannis, D., 2006a. On the quest for chaotic attractors in hydrological processes. *Hydrological Sciences Journal* 51, 1065–1091.
- Koutsoyiannis, D., 2006b. An entropic-stochastic representation of rainfall intermittency: the origin of clustering and persistence. *Water Resources Research* 42.
- Koutsoyiannis, D., 2010. A random walk on water. *Hydrology and Earth System Sciences* 14, 585–601.
- Koutsoyiannis, D., 2000. A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. *Water Resource Research* 36, 1519–1533.
- Koutsoyiannis, D., Montanari, A., 2007. Statistical analysis of hydroclimatic time series: Uncertainty and insights. *Water Resources Research*, 43 (5), W05429, doi:10.1029/2006WR005592, 2007.
- Koutsoyiannis, D., Yao, H., Georgakakos, A., 2008. Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods. *Hydrological Sciences Journal* 53 (1), 142–164.
- Koutsoyiannis, D., Montanari, A., Lins, H.F., Cohn, T.A., 2009. Climate, hydrology and freshwater: towards an interactive incorporation of hydrological experience into climate research—discussion of “The implications of projected climate change for freshwater resources and their management”. *Hydrological Sciences Journal* 54 (2), 394–405.
- Kumar, P., Foufoula-Georgiou, E., 1997. Wavelet analysis for geophysical applications. *Reviews of Geophysics* 35, 385–412.
- Leland, W.E., Taqqu, M.S., Willinger, W., Wilson, D.V., 2002. On the self-similar nature of Ethernet traffic (extended version). *Networking, IEEE/ACM Transactions on* 2, 1–15.
- Mandelbrot, B.B., Van Ness, J.W., 1968. Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10, 422–437.
- Mandelbrot, B.B., 1971. A fast fractional Gaussian noise generator. *Water Resource Research* 7, 543–553.
- Mandelbrot, B.B., Wallis, J.R., 1969. Computer experiments with fractional Gaussian noises: part 1, averages and variances. *Water Resource Research* 5, 228.
- Montanari, A., Rosso, R., Taggu, M., 1997. Fractionally differenced ARIMA models applied to hydrologic time series: identification, estimation, and simulation. *Water Resources Research* 33, 1035–1044.
- Mudelsee, M., 2007. Long memory of rivers from spatial aggregation. *Water Resources Research* 43, W01202.
- O'Connell, P.E., 1974. A simple stochastic modelling of Hurst's law. In: *Proc. International Symposium on Mathematical Models in Hydrology, Warsaw (1971)*, IAHS Publ. No. 100, 169–187.

- Papalexiou, S., 2007. Stochastic modelling of skewed data exhibiting long-range dependence. In: XXIV General Assembly of the International Union of Geodesy and Geophysics.
- Puente, C.E., Sivakumar, B., 2007. Modeling geophysical complexity: a case for geometric determinism. *Hydrology and Earth System Sciences* 11, 721–724.
- Schertzer, D., Tchguirinskaia, I., Lovejoy, S., Hubert, P., Bendjoudi, H., Larchvêque, M., 2002. Which chaos in the rainfall–runoff process. *Hydrological Science Journal* 47, 139–148.
- Shannon Claude, E., Weaver, W., 1948. The mathematical theory of communication. *Bell System Technical Journal* 27, 379–423.
- Sivakumar, B., 2000. Chaos theory in hydrology: important issues and interpretations. *Journal of Hydrology* 227, 1–20.
- Taqqu, M.S., Teverovsky, V., 1998. On estimating the intensity of long-range dependence in finite and infinite variance time series. *A Practical Guide to Heavy Tails: Statistical Techniques and Applications* 177, 218.
- Taqqu, M.S., Teverovsky, V., Willinger, W., 1995. Estimators for long range dependence. *Fractals* 3, 785–798.
- Tyralis, H., Koutsoyiannis, D., 2011. Simultaneous estimation of the parameters of the Hurst-Kolmogorov stochastic process. *Stochastic Environmental Research & Risk Assessment* 25 (1), 21–33.