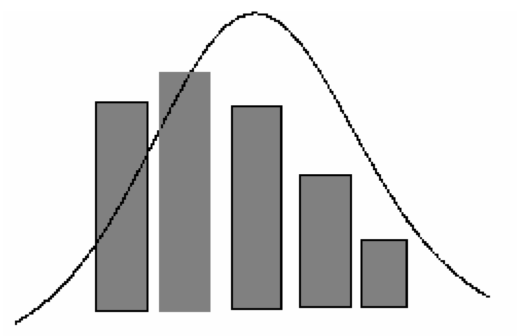
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ON THE STATIONARITY OF RAINFALL TIME SERIES

Carlo De Michele and Harry Pavlopoulos

Department of Statistics
Athens University of Economics and Business

Technical Report No 232, November 2007



DEPARTMENT OF STATISTICS
TECHNICAL REPORT

"ON THE STATIONARITY OF RAINFALL TIME SERIES"

By

C. De Michele*, H. Pavlopoulos**

*DIIAR, Politecnico di Milano, Milano, Italy

**Department of Statistics, Athens University of Economics and Business, Athens, Greece

30 NOVEMBER 2007

0. Abstract

[0] In the analysis of rainfall time series the assumption of s*tationarity* is generally *postulated*. This key assumption undoubtedly facilitates mathematical tractability of stochastic models employed for either descriptive or predictive purposes. However, nature is complex enough so as to accommodate alternations between stationary and non-stationary modes, possibly co-existing in a time series of rainfall measurements. Formal statistical testing of the hypothesis of stationarity in data is very important, since the presence of non-stationarities can lead to seriously questionable results and misinterpretations of data analyses relying on stationarity *ad hoc*.

The present work tests the hypothesis of *weak stationarity* on long rainfall time series records of high resolution (8 years with resolution of 1 hour at Marghera, Italy and 25 years with resolution 5 minutes at Firenze, Italy). This task is carried out with the implementation of non-parametric statistical procedures appropriate for testing homogeneity of mean, variance, and covariance structure, over time. The present work is also concerned with the effects of non-stationarity on the *variance-time function* of series of aggregates over a range of temporal scales of aggregation. Such effects are discussed with regard to observed rainfall data, and with regard to synthetic data obtained from simulations of certain non-stationary model processes constructed here for illustrative purposes.

INDEX TERMS: 1854 (3354) Precipitation, 1869 (Stochastic processes)
KEYWORDS: Non-stationarity, rainfall, scaling properties, tests of stationarity, variance-time plot.

1. Introduction

[1] Stationarity is the most common albeit fundamental assumption made in hydrology and in geophysical sciences at large [*Bras and Rodriguez-Iturbe*, 1985]. However nature's complexity often cannot be represented through a single stationary model. Different types of non-stationarity may be evident in the analysis of natural time series [*Kottegoda*, 1985]. Examples are cyclical changes such as diurnal periodicities acting at time scales smaller than daily but larger than hourly, along with seasonal or quasi-seasonal periodicities acting at weekly up to monthly scales, possibly combined with inter-annual and inter-decadal

cycles and trend effects, due to various sources of non-stationarity attributed to climatic changes or other geophysical forcings, and possibly to anthropogenic impacts on the environment.

[2] The temporal resolution or *sampling frequency* at which a phenomenon is observed affects the evidence upon which the underlying process may be treated as stationary or not. In other words, stationarity is a *scale-dependent* feature of the observed process, in the sense that at some scale of observation (resolution) the hypothesis of stationarity may be accepted, while at some other scale associated with coarser or finer resolution it may be rejected. *Kottegoda* [1985] showed evidence of non-stationarity in *annual series* using the evolutionary spectrum. *Whitcher et al.* [2002] tested stationarity in long *annual series* (Nile river minimum water level) using wavelets. *Wilby* [1997] presented evidence of non-stationarity in historical records of *daily precipitation series* from Britain, attributing it to certain "links" of daily precipitation with several indices of atmospheric circulation dynamics.

[3] Scale-dependence of stationarity pertains also to time series of aggregates with respect to different temporal scales of aggregation. A *series of aggregates* is obtained from a given time series with *fixed* resolution, called the *aggregated series*, by summation of observed values over successive disjoint sections of some given length of interest called *scale of aggregation* (an integer multiple of the resolution of the aggregated series). If the aggregated series is (weakly) stationary, then all series of aggregates inherit (weak) stationarity at every scale of aggregation. If the aggregated series is non-stationary, it is still possible that for large enough scales of aggregation (yet short enough so as to allow sufficient number of sections), some sources of non-stationarity are alleviated, so that the series of aggregates may render stationarity a reasonable assumption to adopt for subsequent analysis or modelling purposes at those scales.

[4] In particular, rainfall time series are characterized by *intermittence* between long periods of *quiescence* with very "light" or no rain, and shorter periods of *intense activity* with very high rain rates, at all scales of observation. Intermittence complicates significantly the assessment of stationarity in time series of rain rate measurements with high resolution. However, by aggregation of rain rate measurements one may obtain series of rainfall accumulation over different temporal scales of aggregation, where it is evident that intermittence becomes less pronounced as the scale of aggregation increases (say from hourly to daily or weekly scale). Therefore, it is anticipated that assessment of stationarity becomes more feasible at larger scales of aggregation, rendering stationarity scale-dependent in this sense too.

[5] Under the assumption of stationarity, *Marani* [2003] argues mathematically on theoretical ground that by aggregation of a stationary stochastic process representing instantaneous rainfall intensity in continuous time, the variance of the resulting stationary process of rainfall aggregates displays several regimes of behavior as a function of the scale of aggregation. These consist of an *inner scaling regime* over a range of small scales of aggregation, an *outer scaling regime* over a range of large scales, and a *transitional regime* over an intermediate range of scales separating the two scaling regimes. The variance is well approximated by a quadratic function of scale (i.e. power-law with exponent 2) over the inner regime, and becomes a power-law function with exponent in the interval [1,2) over the outer regime. The characteristics of the transitional regime and the exact value of the exponent in the outer scaling regime depend upon certain types of asymptotic behavior of the autocorrelation function of the underlying process of rainfall intensity in continuous time, to which the author refers as *memory* and discerns it to finite or infinite memory. These are interesting findings in the sense that from the transitional behavior of variance towards the outer scaling regime, as the scale of aggregation increases, useful information may be drawn. Such information regards to identifying the range of scales in which scaling of variance by a single power-law function of scale holds for series of rainfall aggregates, and also furnishes memory properties (i.e. asymptotic behavior of the correlation structure) of the underlying process being aggregated, under the assumption of stationarity.

[6] To support these theoretical findings, *Marani* [2003] analyzed four long records of rainfall measurements from locations of diverse climates (i.e. Marghera-Italy, Ashover-U.K., Matilija Dam-California, and Lebanon Waterworks-Indiana). However, the very important assumption of stationarity was not verified in any quantitative way on the data. Instead, stationarity was merely assumed to hold ad hoc, during any given month of the year (e.g. every January), across all years (i.e. interannually) in the record available at each site. Therefore, while the issue of testing stationarity withstands the analysis of data presented therein, the empirical evidence of two power-law scaling regimes separated by a transitional regime in variance-time plots, in our opinion may be merely an artifact of non-stationarities present in the analyzed data.

[7] The purpose of the present work is twofold. First, to highlight the possibility of non-stationarity in rainfall time series of high temporal resolution, by raising sufficient statistical evidence against the hypothesis of stationarity. To this end, some formal statistical procedures suitable for testing the hypothesis of (weak) stationarity are presented in Section 2. Results from the application of these statistical procedures

on long rainfall records of high resolution, including the series from Marghera considered by *Marani* [2003], are presented and discussed in Section 3. The second goal is to provide some understanding of the effects of non-stationarity on the behavior of variance-time plots of series of aggregates obtained from non-stationary processes. This task is undertaken in Section 4 with the construction of suitable examples of non-stationary processes for illustrative purposes. Synthetic time series from these non-stationary models do demonstrate double scaling regime structure in their variance-time plots. Our conclusions are summarized in Section 5.

2. Weak stationarity and statistical testing procedures

2.1 Notation and the concept of weak stationarity

[8] Let $\{X_t, t \in N\}$ denote a time series process observed at discrete time instants $t \in N$, where the index set N denotes the set of natural numbers. Introducing notation and terminology, $M_X(t) = E\{X_t\}$ is the *mean* function, $C_X(t,s) = Cov(X_t, X_s) = E\{[X_t - \mu_X(t)]\cdot[X_s - \mu_X(s)]\}$ is the *covariance* function, and $\sigma^2_X(t) = C_X(t,t) = E\{[X_t - \mu_X(t)]^2\} = Var\{X_t\}$ is the *variance* function of the process, which is said to be a *weakly stationary* process or *wide sense stationary* or *second-order stationary* if and only if the following three requirements are fulfilled:

($S0$) $E[X_t^2] < \infty$, for all $t \in N$,

($S1$) $M_X(t) = \mu_X$, for all $t \in N$,

($S2$) $C_X(t, s) = C_X(t + \tau, s + \tau)$, for all t, s, $\tau \in N$.

The above definition requires that all moments up to second order are finite, and in particular that the mean function remains constant, and that the covariance function remains invariant under translations of the origin of time, or equivalently that it reduces to an *even* and *non-negative definite* function γ_X of temporal lag only,

$$\gamma_X(|t-s|) = C_X(t,s) = C_X(s,t) = \begin{cases} C_X(0,s-t), & t \le s \\ C_X(t-s,0), & t \ge s \end{cases},$$

whence the variance function becomes also constant, i.e. $\sigma^2_X(t) = C_X(t,t) = C_X(0,0) = \gamma_X(0) = \sigma^2_X$.

[9] Weak stationarity as defined above is the most common and broad sense in which the term "stationarity" is usually interpreted or considered, allowing also several other stronger notions of stationarity to hold along with it [e.g. see *Priestley*,1981-a; *Yaglom*,1987; *Brockwell and Davis*,1991]. In general, stationarity is a concept addressing stability or homogeneity of the probabilistic behavior of a random process over time. Such stability depends much on the physical mechanisms generating the phenomenon measured by the random process, and as noted in the Introduction it may depend on the scale at which the random process is observed. Some interesting viewpoints on the concept of stationarity have been expressed by *Mandelbrot* [1983, pp. 383-386].

[10] Non-stationarity of a process may be inferred from significant changes over time in either the mean, or the variance, or the covariance of the process. On the contrary, weak stationarity is a safe assumption to make only if both mean and variance of the process remain constant, and also the covariance function does not vary over time, with respect to some quantitative criteria of confidence. For example, given a non-stationary process $\{X_t\}$ with finite mean function $\mu_X(t)$, finite and positive variance function $\sigma_X^2(t)$, and covariance function $C_X(t,s)$, it is clear that the process $\{Y_t = [X_t - \mu_X(t)]/\sigma_X(t)\}$ has constant mean $\mu_Y(t) \equiv 0$, constant variance $\sigma_Y^2(t) \equiv 1$, but its covariance function $C_Y(t,s) = C_X(t,s)/[\sigma_X(t)\cdot\sigma_X(s)]$ need not be invariant through time, unless if special conditions are imposed on the X-process. Thus, the *Y*-process is anticipated to have inhomogeneous covariance function, in general, and therefore to be non-stationary although its mean and variance are constant.

[11] Several statistical procedures are available in the literature for testing the hypothesis of weak stationarity on actual time series data of fixed resolution, regarding different aspects of the hypothesis of weak stationarity, namely homogeneity of mean, homogeneity of variance, and homogeneity of covariance. Existence of moments up to second order, which is part of the definition of weak stationarity, is an inescapable assumption made *a priori* in all these tests. Since no further assumptions are made here about marginal or bivariate probability distributions of the observed process (except existence of moments up to second order), our interest is focused only on non-parametric testing procedures, briefly presented in the rest of this section and applied to actual rainfall data in Section 3.

2.2 Testing homogeneity of mean

[12] Here we propose a new and relatively simple procedure that we used for testing homogeneity of mean. Given a time series process $X=\{X_t,\ t \in N\}$ and a temporal scale of aggregation $m \geq 1$, the time series process $X^{(m)}=\{X_t^{(m)}, t \in N\}$ of *m-aggregates* of X is defined by taking $X_t^{(m)} = \sum_{i=(t-1)m+1}^{tm} X_i$. Clearly, if the process X has constant mean μ_X , then the process $X^{(m)}$ also has constant mean given by the formula

$$\mu_{X^{(m)}} = m \cdot \mu_X . \tag{1}$$

In fact, if X is weakly stationary, then $X^{(m)}$ is also weakly stationary, with covariance function explicitly determined by covariances of the original X-process (e.g. see *Koutsoyiannis* [2002]). In general, we shall refer to the functional relationship between the mean of m-aggregates and the scale of aggregation m, as *mean-time function*, and to the plot of sample mean of m-aggregates versus the scale m as *mean-time plot*. In the case of homogeneous mean, equation (1) represents a linear mean-time function, and the mean-time plot of the sample mean $\hat{\mu}_{X^{(m)}}$ of m-aggregates against the scale of aggregation m is expected to fit well along a straight line. The intercept term of the best fitted line (say by the method of ordinary least squares) ought to be statistically insignificant, and its slope can be used as a working estimate $\hat{\mu}_X$ of the constant mean μ_X .

[13] Now suppose that a long time series $\{x_t\}_{t=1}^{T}$ is divided into $n \geq 2$ sub-series over disjoint large sections of equal length. In the case of homogeneous mean μ_X , the mean-time plot of every subseries (over the same scales of aggregation implemented in every section) should yield a strong linear fit. As already pointed out, all intercepts of the fitted lines (one line for each subseries or section) ought to be statistically insignificant, and the slopes of the fitted lines provide an ensemble of sample estimates $\{\hat{\mu}_{X_j} ; j = 1,\cdots,n\}$ of the true constant mean μ_X . Such estimates facilitate the construction of a two-sided confidence interval for μ_X , on the basis of which one may assess further the statistical significance of the hypothesized homogeneity of the mean, at desired level of significance $0 < \alpha < 1$, rejecting the hypothesis of homogeneity of the mean if the percentage of the n estimates $\{\hat{\mu}_{X_j} ; j = 1,\cdots,n\}$ falling outside that interval is greater than α. Alternatively, homogeneity of the mean may also be rejected if the overall sample mean $\bar{x} = (1/T) \cdot \sum_{t=1}^{T} x_t$, or if the slope $\hat{\mu}_X$ of the overall mean-time plot, falls outside that interval.

[14] For example, if the estimates $\{\hat{\mu}_{X_j} ; j = 1,\cdots,n\}$ are not correlated and the shape of their distribution is not too skewed, an approximate confidence interval for μ_X (for sufficiently large n) is obtained by the formula

$$\overline{\mu}_X \pm z_{1-\alpha/2} \cdot s(\overline{\mu}_X), \qquad (2)$$

where $\overline{\mu}_X = n^{-1} \cdot \sum_{j=1}^{n} \hat{\mu}_{X_j}$ is a new sample estimate of the true constant mean μ_X, $z_{1-\alpha/2}$ is the (1-α /2)-quantile of the standard normal distribution, and $s(\overline{\mu}_X) = \sqrt{[n \cdot (n-1)]^{-1} \cdot \sum_{j=1}^{n} \left(\hat{\mu}_{X_j} - \overline{\mu}_X\right)^2}$ is the standard error of the estimate $\overline{\mu}_X$ (e.g. see *Gilbert* [1987, sections 11.7, 11.12]).

[15] If the hypothesis of a constant mean is not rejected, then one should proceed to test for homogeneity of variance and covariance in order to establish validity of the assumption of stationarity. Otherwise, if there is sufficiently significant evidence pointing to rejection of homogeneity of the mean, then the process should be considered non-stationary *per se*. However, it is worthy to remark that, if one suspects that inhomogeneity of the mean is not the only source of non-stationarity, then one ought to estimate (parametrically or non-parametrically) and remove the estimated inhomogeneous mean function (i.e. trend) from the initial series of observed data. Subsequently, the *de-trended* series should be subjected to tests of homogeneity of variance and covariance, and possibly to a test of homogeneity of mean anew as an additional check on the quality of the de-trending procedure.

2.3 Testing homogeneity of variance

[16] *Pagan and Schwert* [1990] proposed three non-parametric statistical procedures for testing homogeneity of variance over time, on the basis of regularly sampled time series data. Namely, these are the *post-sample prediction* test, the *cumulative sum* test (CuSum in short), and the *modified scaled-range* test, and applied them to investigations of weak stationarity in stock market financial data. The CuSum test is closest to the original idea proposed by *Mandelbrot* [1963] for exploration of existence of variance, and homoscedasticity thereof, formalizing that idea into a rigorous testing procedure, which is presented here briefly in order to be implemented in Section 3 for the detection of heteroscedasticity in rainfall time series.

[17] The CuSum test presumes that the observed data $\{y_t\}_{t=1}^{T}$ are generated by a process $\{Y_t; t \in N\}$ with mean zero, and examines the variability of the cumulative sums of the series of squared data $\{y_t^2\}_{t=1}^{T}$, up to *r*-fractions of the time of observation *T*, through the normalized statistic given by the formula

$$\varphi(r) = \frac{1}{\sqrt{\hat{\nu} \cdot T}} \cdot \sum_{t=1}^{[r \cdot T]} \left(y_t^2 - \hat{\sigma}_T^2 \right), \tag{3}$$

where $0 < r < 1$ denotes fraction of time, [.] is the integer part function, $\hat{\sigma}_T^2 = (1/T) \cdot \sum_{t=1}^{T} y_t^2$ is a sample estimate of variance over the entire record of data (of length *T*), and $\hat{\nu} = \hat{\gamma}_0 + 2 \cdot \sum_{k=1}^{8} \hat{\gamma}_k \cdot \left(1 - \frac{k}{9}\right)$, with $\hat{\gamma}_k$ being the sample covariance estimate (at lag $k \geq 0$) computed from the whole series of the squared data $\left\{ y_t^2 \right\}_{t=1}^{T}$. Under the hypothesis of weak stationarity, $\hat{\nu}$ serves as an estimate of $\nu = \gamma_0 + 2 \cdot \sum_{k=1}^{\infty} \gamma_k$, with γ_k being covariance (at lag $k \geq 0$) of the squared process $\left\{ Y_t^2 ; t \in N \right\}$, provided that ν is a finite number, i.e. if the squared process $\left\{ Y_t^2 ; t \in N \right\}$ possesses finite moments up to second order and also has finite memory (see *Phillips* [1987] for explicit conditions).

[18] The asymptotic probability distribution of the statistic $\varphi(r)$, for sufficiently large *T*, is a normal distribution with mean 0 and variance $r(1-r)$, a result derived by *Lo* [1988] under a set of suitable conditions formulated by *Phillips* [1987]. This asymptotic result facilitates the construction of a non-parametric confidence region, rejecting the hypothesis of homoscedasticity at desired level of significance $0 < \alpha < 1$ if the plot of $\varphi(r)$ wanders-off that region during a gross fraction of *T* that is greater than α. The boundary of the confidence region is a closed curve comprising of two arcs, $c_\alpha^+(r) = z_{1-\alpha/2} \cdot \sqrt{r(1-r)}$ and $c_\alpha^-(r) = z_{\alpha/2} \cdot \sqrt{r(1-r)} = -c_\alpha^+(r)$, meeting at the beginning ($r = 0$) and at the end ($r = 1$) of the time of observation (i.e. $c_\alpha^-(0) = c_\alpha^+(0) = 0 = c_\alpha^-(1) = c_\alpha^+(1)$).

2.4 Testing homogeneity of covariance

[19] The tests developed by *Pagan and Schwert* [1990], may also be considered as tests of homogeneity of covariance structure, albeit only implicitly as such, in the sense that while testing directly for homoscedasticity, under the null hypothesis, the asymptotic behavior of the test statistic $\varphi(r)$ presumes stationarity of the covariance structure as well. However, lack of evidence towards rejection of homoscedasticity tested by the procedures proposed by *Pagan and Schwert* [1990], may be insufficient information for acceptance of the stronger hypothesis of stationary covariance.

[20] Spectral analysis in the frequency domain is generally a more appropriate framework for drawing useful information about stationarity or non-stationarity of covariance structure. For example, it is well known that for *scale-invariant* processes the power spectrum $S(\omega)$ is a power-law function of frequencies ω within a certain range of temporal scales [*Mandelbrot*, 1983; *Davis et al.*, 1996; *Kärner*, 2002]. That is, for $\omega \in (\omega_l, \omega_u)$, there is $\delta \geq 0$, so that $S(\omega)$ is proportional to $\omega^{-\delta}$, and the spectral exponent δ is informative about stationarity or non-stationarity of the underlying scale-invariant process. Specifically, the Wiener-Khinchine theorem implies that values of $\delta < 1$ correspond to stationary processes, and the value $\delta = 0$ is associated with white noise. In contrast, values in the range $1 < \delta < 3$ are associated with non-stationary processes having stationary increments [*Mandelbrot*, 1998, pp.74-79; *Malamud and Turcotte*, 1999].

[21] Several studies on observed rainfall series support the power-law type of behavior of the spectrum, with exponents near zero ($\delta \approx 0$) for time scales longer than 1-2 weeks, $0 < \delta < 1$ in the range of scales from 1-2 weeks down to 1-6 hours, and $1 < \delta < 2$ for time scales smaller than 1-6 hours [e.g. see *Fraedrich and Larnder*, 1993; *Georgakakos et al.*, 1994; *Fabry*, 1996; *Harris et al.*, 1996; *Tessier et al.*, 1996; *Olsson and Burlando*, 2002]. These studies verify scale-dependence of stationary and non-stationary behavior in rainfall time series, where stationarity tends to settle over a range of larger scales of observation or aggregarion, and non-stationarity is more prominent across a range of smaller scales. Similar behavior has also been observed in time series of other geophysical variables [e.g. see *Davis et al.*, 1996; *Kärner*, 2002].

[22] The basic tool of spectral analysis is the *discrete (or finite) Fourier transform* of the data, whose squared modulus is known as *periodogram*, and through which various estimators of the spectrum are constructed and studied in the literature. Statistical properties of periodogram based estimators of power spectra are well established for weakly stationary processes (e.g. see *Priestley* [1981-a], *Brockwell and Davis* [1991]). However, when one is concerned about the very assumption of weak stationarity, then the concept of *evolutionary spectrum* introduced by *Priestley* [1965] is naturally more appealing and appropriate for the study of inhomogeneities in the covariance structure of the underlying process (see also *Priestley* [1981-b]).

[23] An important advancement towards rigorous statistical testing of the hypothesis of weak stationarity has recently been made by *Ahamada and Boutahar* [2002], with the construction of a non-parametric procedure based on the concept of evolutionary spectrum. The main advantage of this new test, compared to previous ones [*Priestley and Subba Rao*, 1969; *Pagan and Schwert*, 1990], is that it is highly sensitive to a great variety of types of instabilities in the covariance structure, and it is not confined by design to detect particular types of instability associated with certain kinds of non-stationarity. This high sensitivity is due to the local character of the Fourier transform of the data, subsequently built into the estimator of the evolutionary spectral density and also into the proposed test statistic. Here we briefly describe the stationarity test of *Ahamada and Boutahar* [2002], in order to apply it on time series of rainfall in Section 3.

[24] Consider a time series process $\{Y_t\}$ with zero mean, defined as a stochastic integral

$$Y_t = \int_{-\pi}^{+\pi} A_t(\omega) e^{i\omega t} d\zeta(\omega) \tag{4}$$

with respect to a random ζ-process possessing orthogonal increments $d\zeta(\omega)$ on $[-\pi, \pi]$, such that $E[d\zeta(\omega)] = 0$ and $E[|d\zeta(\omega)|^2] = \mathrm{d}\mu(\omega)$, where $\mu(\omega)$ is a measure on $[-\pi, \pi]$. The integrated (random) function $A_t(\omega)$ is defined at frequencies $\omega \in [-\pi, \pi]$, and for each fixed $\omega \in [-\pi, \pi]$ the modulus of the (generalized) Fourier transform of the t-sequence $\{A_t(\omega)\}$ is supposed to have an absolute maximum at the origin (i.e. at $\omega = 0$). Then, the evolutionary spectral density $h_t(\omega)$ of such an oscillatory process $\{Y_t\}$ is defined as

$$h_t(\omega) = \frac{dH_t(\omega)}{d\omega}, \text{ for } \omega \in [-\pi, \pi], \tag{5}$$

where $dH_t(\omega) = |A_t(\omega)|^2 d\mu(\omega)$. An estimator $\hat{h}_t(\omega)$ of the evolutionary spectral density $h_t(\omega)$ is

$$\hat{h}_t(\omega) = \sum_{v \in Z} w_v |U_{t-v}(\omega)|^2 , \tag{6}$$

where $U_t(\omega) = \sum_{u \in Z} g_u y_{t-u} e^{-i\omega(t-u)}$, g_u is a Bartlett window, and w_v is a Daniel window, defined respectively by

$$g_u = \begin{cases} 1/\left(2\sqrt{h\pi}\right) & \text{if } |u| \le h \\ 0 & \text{if } |u| > h \end{cases} \quad \text{and} \quad w_v = \begin{cases} 1/T' & \text{if } |v| \le T'/2 \\ 0 & \text{if } |v| > T'/2 \end{cases}. \tag{7}$$

[25] Given time series data $\{y_t\}_{t=1}^{T}$ from a process $\{Y_t\}$ in the class considered above, we take $h = 7$, $T' = 20$, and consider the set of times $\{t_i = 20i\}_{i=1}^{I}$, where $I = [T/20]$, and the set of frequencies $\{\omega_j = (\pi / 20) \cdot (1 + 3(j-1))\}_{j=1}^{7}$, as suggested by *Ahamada and Boutahar* [2002], so that certain desired properties are met by the estimates $\hat{h}_{t_i}(\omega_j)$. Then, set $V_{ij} = \log \hat{h}_{t_i}(\omega_j)$, $V_i = (1/7) \cdot \sum_{j=1}^{7} V_{ij}$, $V_j = (1/I) \cdot \sum_{i=1}^{I} V_{ij}$, $\hat{\mu}_T = (1/(7 \cdot I)) \cdot \sum_{i=1}^{I} \sum_{j=1}^{7} V_{ij}$, $\hat{\sigma}^2 = (1/I) \cdot \sum_{i=1}^{I} (V_i - \hat{\mu}_T)^2$, and $S_r = \left(1/\left(\hat{\sigma}\sqrt{I}\right)\right) \cdot \sum_{i=1}^{r} (V_i - \hat{\mu}_T)$, for $r = 1,\ldots, I$, for $i = 1,...,I,$ and for $j = 1,\ldots, 7$.

[26] Under the null hypothesis of weak stationarity of the process $\{Y_t\}$, the limit probability distribution of the statistic $S = sup\{|S_1|, \cdots, |S_I|\}$, for T large enough, has cumulative distribution function given by

$$F_S(x) = 1 - 2 \cdot \sum_{k=1}^{\infty} (-1)^{1+k} exp(-2k^2 x^2). \tag{8}$$

Thus, a test of the null hypothesis of covariance stationarity obtains, at desired level of significance $0 < \alpha < 1$, rejecting the hypothesis if the observed value of the statistic S is greater than the (1-α) quantile of its limit distribution.

Closing this section we should remark that proper implementation of either the CuSum test or of the evolutionary spectrum test on non-negative time series of rainfall data $\{x_t\}_{t=1}^{T}$ of positive mean, requires that the test is applied to the "de-meaned" series $\{y_t = x_t - \bar{x}\}_{t=1}^{T}$, provided that the series $\{x_t\}_{t=1}^{T}$ has passed a test of homogeneity of mean, and in that case its mean is estimated by the sample mean $\bar{x} = (1/T) \cdot \sum_{t=1}^{T} x_t$.

3. Testing weak stationarity on rainfall data

[27] The tests presented in Section 2 are applied here in order to check temporal homogeneity of mean, variance and covariance on two datasets of rainfall observations, one collected at Marghera over the eight-year period 1993-2000 with resolution *1 hour*, and the other at Firenze, Ximeniano Observatory over the twenty-five-year period 1962-1986 with resolution *5 minutes*. The first record is available at the web site http://www.istitutoveneto.it/venezia/dati/atmosfera, and it is one of the four long records analyzed by *Marani*

[2003]. Respecting the monthly segmentation of each single year in these two time series, the tests were separately performed on each month of every year.

[28] Initially the mean-time plot was calculated over each of the 12 months in each year of the Marghera and Firenze data. The range of scales of aggregation implemented on each month of Marghera data is from 1 hour to 93 hours (with all the integer submultiples of the total number of data (744): 2, 3, 4, 6, 8, 12, 24, 31, 62 hrs), and the range of scales of aggregation implemented on each month of Firenze data is from 5 minutes to 93 hours (with all the integer submultiples of the total number of data, 8928). All the obtained mean-time plots agree perfectly with the anticipated linear behavior represented by equation (1). Coefficients of determination are strikingly high (exceeding 0.99 in all cases), intercept terms are insignificant indeed, residuals in each regression are non-correlated and their standard error is markedly low. The slope of each monthly mean-time plot at Marghera (obtained via simple linear regression by ordinary least squares) is reported in Table 1. Similar information was obtained for Firenze (not reported here). These findings do point to homogeneity of mean over each individual monthly section in both Marghera and Firenze.

[29] However the estimated slopes of fitted lines vary from month to month within any given year, due to anticipated seasonal effects, but also vary significantly from year to year for every given month of the year. Using formula (2), a confidence interval at level of significance $\alpha = 0.05$ was calculated for each month, based on the ensemble of estimated slopes obtained for that month from all the available years, in order to assess the significance of variability of monthly mean from year to year. For Marghera, these confidence intervals are reported in Table 1, and highlighted gray slots indicate months where the slope estimate from the corresponding mean-time plot lies outside the 95% confidence interval. Figures 1 and 2 depict these 12 intervals for Marghera and Firenze respectively. Note that confidence intervals for Firenze are considerably "tighter" than those calculated for Marghera, a feature explained by the fact that the number of observed years in Firenze is nearly triple that of Marghera (25 versus 8 years) and the resolution in Firenze is twelve times higher than that of Marghera (5 minutes versus 1 hour). Clearly, for every given month in either figure, the percentage of estimated slopes falling outside the corresponding confidence interval is by far greater than 5%. Thus, we may claim with 95% confidence that the interannual variability of monthly mean of rainfall, over any given month of the year, deviates significantly from the hypothesis of homogeneity, at both Firenze and Marghera.

[30] Evidence of homogeneity of the mean during every individual month of every probed year in both Marghera and Firenze, on the basis of linear regression statistics obtained from mean-time plots, allows direct implementation of CuSum-test and evolutionary spectral test on each monthly section, after subtraction of the corresponding sample mean from the observations of that month. Tables 2 and 3 report *p*-values obtained from the implementation of CuSum test (for each month of each year) on Marghera and Firenze data respectively. Similarly, Tables 4 and 5 report *p*-values obtained from the evolutionary spectrum test applied on the same monthly sections of Marghera and Firenze data too.

[31] The symbol *NaN* appearing in some slots of Tables 3 and 5, regarding mostly summer months at Firenze, indicates that the observed rainfall was constantly zero (i.e. no rain) during those months of those particular years, thus supporting the hypothesis of stationarity despite the fact that a *p*-value is not obtainable in those cases. Highlighted gray slots in Tables 2-5 indicate months where the corresponding *p*-values are below the 5% level of significance, and therefore the hypothesis of stationarity ought to be rejected in each of those months with 95% confidence. For example, the month of January at Marghera supports the hypothesis of homoscedasticity at 5% level of significance in four out of the eight observed years (see Table 2), but supports the hypothesis of homogeneous covariance (also at 5% level) only in two out of the eight years (see Table 4), incidentally a subset of the four slots where homoscedasticity is supported. Figure 3 shows the variability of the φ-statistic given by (3), during the month of January at Marghera in the year 1997 (figure 3a) and in the year 1998 (figure 3b), and the rejection region associated with the CuSum test at 5% level of significance, respectively.

[32] The CuSum test and the covariance stationarity test were also applied on the annual sections of the two time series, and to the entire record of each of these two series, and *p*-values of the tests are also reported in Tables 2-5. The hypothesis of homoscedasticity is rejectable by the CuSum test (at 5% level), in 5 out of the 8 available years for Marghera (percentage 63%), and in 17 out of 25 available years for Firenze (percentage 68%). The hypothesis of homogeneous covariance is rejectable by the evolutionary spectrum test (at 5% level), in 6 out of 8 years for Marghera (percentage 75%), and in all 25 years (percentage 100%) for Firenze. Both tests reject the hypotheses of weak stationarity (at 5% level) in both Marghera and Firenze, when applied to the entire records of available data.

[33] Overall, we see that in either monthly, or annual, or global time scales, the information from Firenze supports the rejection of weak stationarity more decisively than the information from Marghera, which may

be a consequence of the significant differences in the length and the temporal resolution of the two records, besides other interpretations of climatological content that one might speculate. In any event, the analysis just presented raises rigorous statistical evidence against the hypothesis of weak stationarity of observed rainfall processes at Marghera and Firenze. In light of this evidence, the assumption of weak stationarity is rejectable for the Marghera data, whose variance-time plot was studied by *Marani* [2003] assuming stationarity.

[34] Figures 4 and 5 show evidence of scale-dependence of stationarity in the two records of data from Marghera and Firenze respectively. Specifically, a raw periodogram estimate of the power spectrum was obtained for the month of January in every year, and the average of those estimates (8 for Marghera and 25 for Firenze) is plotted against (Fourier) frequencies (as log-log-plot). Three bands of frequencies have been identified in each of the two plots, each band expressing a power-law function of frequency. For Marghera, there is a band of high frequencies corresponding to small temporal scales up to 1.5 hours, where the estimated value 1.97 of the spectral exponent δ is clearly greater than unity, a band of intermediate frequencies corresponding to time scales ranging from 1.5 hours to 120 hours (about 1 week), where the spectral exponent shifts to a value close to unity (estimate 1.06), and a band corresponding to larger scales ranging from 120 hours to one month, where the spectral exponent is nearly zero, thus indicating white noise (see Figure 4). For Firenze, three different bands of frequencies are shown in Figure 5, and the corresponding estimates of spectral exponent δ are, 1.19 in the range from 5 minutes to 1 hour, 0.77 over the intermediate range of time scales from 1 hour to 1 day, and 0.22 (i.e. closer to zero) for larger scales ranging from 1 day to 1 month.

[35] Although these results, based on raw periodogram estimates, are not as rigorous as those obtained by the formal statistical tests presented above, they are in good agreement with similar results reported by *Olsson and Burlando* [2002], who used four time series of rainfall (Camaldoli, Firenze, Livorno, and Vallombrosa) with a resolution of 20 minutes from year 1962 up to 1986, and found power-law spectra with $\delta > 1$ for time scales shorter than 2-6 hours (depending on the location), and with an exponent $\delta < 1$ for scales ranging from 2-6 hours to around 1 week up to 1 month. Note that our data for Firenze cover the same span of years (1962-1986), but at finer resolution (5 minutes), which affects the bands where the spectrum behaves as a power-law function of frequency, and also affects the estimates of spectral exponents. In any event, these results along with those obtained by *Olsson and Burlando* [2002] confirm non-stationarity over small time scales and a tendency towards stationarity over large time scales, i.e. scale-dependence of stationarity.

Closing this section we should like to mention that, as a check of the performance of the tests presented above, we also applied them to synthetic time series data obtained from simulations of specific stationary model processes with long-range dependence, namely fractional Gaussian noise processes (see Section 4). In all those cases, the hypothesis of stationarity was accepted at the canonical levels of significance 5% and 1%.

4. Effects of non-stationarity on the variance-time plot

4.1 Pertinence to long-range dependence

[36] The functional relationship between the variance of the process $X^{(m)}$ of m-aggregates, of a given process X, and the scale of aggregation m, is known as *variance-time function* of X, and the plot of sample variance of m-aggregates versus the scale m is referred to as *variance-time plot*. The variance-time plot is one among several statistical procedures available in the literature for detecting the presence of long-memory or long-range dependence in a stationary process. A stochastic process possessing this property is referred to as *long-range dependent* process, in short LRD. A thorough monograph about LRD processes and their statistics is that of *Beran* [1994].

[37] The very notion of long-memory remains open to controversy among statisticians and other scientists; see Sections 3.2, 3.7, and 8.3 in *Embrechts and Maejima* [2002]. The most commonly adopted interpretation or definition of the concept of LRD is that of a *stationary process* with covariance function asymptotically proportional to a power-law function at very large lags, or equivalently with spectral density function asymptotically proportional to another power-law function at very low frequencies. In fact, the exponents of these two asymptotic power-laws (in time domain and in frequency domain) are determined or characterized by a common parameter of memory denoted by H, referred to as *Hurst exponent*, and taking values in the open interval $0.5 < H < 1$. By any definition, however, long-memory presumes (weak) stationarity of the LRD process. Thus, rigorous testing of the hypothesis of (weak) stationarity is very much a key issue, to be addressed before any interpretation of results obtained from analysis of the variance-time plot, or of any other tool available for long-memory analysis. Consequently, a relevant issue is to understand how the behavior of the variance-time plot is influenced by various sources of non-stationarity in the process of interest (i.e. the aggregated process). The examples that we consider in this section contribute to this goal.

[38] *Teverovsky and Taqqu* [1997] and *Dang and Molnar* [1999] showed that it is difficult to discern an LRD-process from a non-stationary process. Similar concerns have also emerged in recent econometric literature [*Diebold and Inoue* (2001) and *Mikosch and Stărică* (2000, 2001)] about the subtlety of long-memory, arguing that long-memory is falsely claimed or inferred in some cases of non-stationary Markov processes with short-memory (e.g. GARCH processes), while it is merely an *artifact* or *disguise* of non-stationarity. *Bhattacharya, Gupta, and Waymire* [1983] started the investigation of long-memory (via R/S-analysis) in the presence of some mild deviations from stationarity, namely under the influence of what they refer to as "slow trend"; also see *Künsch* [1986] and *Bhattacharya and Waymire* [1990, Section I-14]. *Dang and Molnar* [1999] studied the influence of two different deterministic sources of non-stationarity (level shifts and linear trends) on long-range dependence, and investigated the associated problem of estimation of the Hurst exponent H relying on various methodologies and tools (i.e. variance-time plot, R/S-analysis, periodogram analysis, and wavelets). Estimates of H based on methods that rely strongly on the assumption of stationarity (i.e. variance-time plot and periodogram analysis) were poor compared to estimates obtained by methods that are more robust to non-stationarities (i.e. R/S-analysis and wavelet based estimation), pointing to a partial conclusion that variance-time plots should not be trusted without preliminary testing of the assumption of stationarity.

4.2 Illustrating examples

[39] The stochastic ingredient of our examples is drawn from the class of *fractional Gaussian noise* (FGN in short) processes, originally introduced by Mandelbrot and his co-workers in an effort to understand the famous *Hurst effect* [e.g. see *Beran*, 1994; *Samorodnitsky and Taqqu*, 1994; *Koutsoyiannis*, 2002]. A process $\{X_t\}$ belongs to the FGN class if and only if it is a stationary Gaussian process, say with mean μ_X and variance σ_X^2, and with covariance function defined on integer lags h by $\gamma_X(h) = (\sigma_X^2 / 2) \cdot \{(|h|+1)^{2H} - 2|h|^{2H} + (|h|-1)^{2H}\}$, in short writing $X \sim FGN(\mu_X, \sigma_X^2, H)$. It can be shown that an FGN process with $0.5 < H < 1$ is an LRD process, whence the parameter H is referred to as *memory parameter* or *Hurst exponent*. Due to the importance of FGN processes in modeling random phenomena with long memory, several computational algorithms are broadly available for their numerical simulation, and herein we have implemented the algorithm of *Paxson* [1995].

[40] Our first example is a process $Z = \{Z_t\}$, such that $Z_t = X_t + A_t$ for every *t*, where $X = \{X_t\}$ is a $FGN(\mu_X, \sigma_X^2, H)$ process with long memory (i.e. $0.5 < H < 1$), and $A_t = a \cdot sin(2\pi \cdot t / T)$ is a deterministic sinusoidal trend function with period *T* and amplitude *a*. The parameter a / σ_X introduces a "signal-to-noise" type of ratio for comparison of the amplitude of the harmonic trend against the standard deviation of the FGN process. The variance of the process $\{X_t^{(m)}\}$, of *m*-aggregates of the *X*-process at given scale of aggregation $m \geq 1$, is given by the formula $\sigma_{X^{(m)}}^2 = m^{2H} \cdot \sigma_X^2$ (e.g. see *Beran*, 1994). An immediate consequence of this formula is that the plot of $\log(\sigma_{X^{(m)}}^2)$ versus $\log(m)$ is a straight line with slope 2*H* and intercept $log(\sigma_X^2)$. That is, the variance-time plot of an FGN process with Hurst exponent $0.5<H<1$, is a power-law function of the scale of aggregation *m*, with (scaling) exponent $1<2H<2$.

[41] Since *m*-aggregates $\{A_t^{(m)}\}$ of the harmonic trend also form a deterministic function of time, and $Z_t^{(m)} = X_t^{(m)} + A_t^{(m)}$ for every *t*, it follows that the processes of aggregates $\{Z_t^{(m)}\}$ and $\{X_t^{(m)}\}$ have the same variance. That is, $\sigma_{Z^{(m)}}^2 = \sigma_{X^{(m)}}^2 = m^{2H} \cdot \sigma_X^2$, for every scale of aggregation $m \geq 1$, showing that the two processes also have exactly the same variance-time function (a single power-law with exponent $2H$), although *Z* and its aggregates are non-stationary processes, while *X* and its aggregates are stationary. In particular, when the scale of aggregation is some integer multiple of the harmonic period *T*, say $m = k \cdot T$, with integer $k \geq 1$, then it is elementary to see that $A_t^{(kT)} = 0$, and therefore $Z_t^{(kT)} = X_t^{(kT)}$, showing that in such case the two processes of aggregates have identical sample paths (not just equal variances). That is, the non-stationarity of aggregates of the *Z*-process is totally eliminated by this particular choice of scale of aggregation.

[42] A synthetic time series of length $n = 2^{14}$ was obtained by simulation of an $FGN(0, 1, 0.7)$ process, whose initial section for $1 \leq t \leq 1000$ is shown in Figure 6(a), and Figure 6(b) depicts the harmonic function A_t with period $T = 50$ and amplitude $a = 1$ over the same section of instants *t*. The ratio *n/T* quantifies the number of complete periods (of the sinusoidal trend) spanned by the simulated time series. Figure 6(c) depicts the initial section of the synthetic time series of the non-stationary *Z*-process, obtained by the superposition $Z_t = X_t + A_t$ of the harmonic trend and the simulated FGN series. Note that $a/\sigma_X = 1$ and $\left[n/T\right] = 327$. From the data of the entire synthetic series of the *Z*-process we computed an estimate of its variance-time plot for scales of aggregation in the range $1 \le m \le 500$. The results are shown in Figure 6(d), in a log-log plot (i.e. $\log\left(\hat{\sigma}^2_{Z^{(m)}}\right)$ vs. $\log(m)$), together with the estimate of the variance-time plot of the synthetic series of the *X*-process (i.e. $\log\left(\hat{\sigma}^2_{X^{(m)}}\right)$ vs. $\log(m)$). The estimated variance-time function from the simulated *X*-series fits quite well the theoretical one with $H = 0.7$ (the fitted exponent is $\hat{H} = 0.688$).

[43] Figure 6(d) shows that the variance-time plots of $\{Z_t\}$ and $\{X_t\}$ tend to coincide for scales of aggregation $m \ge 150$, defining an “outer” regime of power-law scaling of the variance of aggregates. An “inner” regime of power-law scaling of variance is also very distinct in Figure 6(d), over the range of small scales of aggregation $m < 25$. These two scaling regimes are separated by a “transitional” regime over the intermediate range of scales $25 < m < 150$. This transitional regime presents an interesting structure, alternating between “plateaus” (e.g. $25 < m < 50$, $75 < m < 100$, $125 < m < 150$), and intermediate scaling regimes (e.g. $50 < m < 75$, $100 < m < 125$) with exponents approaching the exponent of the outer regime. For example, over the four ranges of scales $1 \le m \le 25$, $50 \le m \le 75$, $100 \le m \le 125$, and $m \ge 150$ we obtained estimates of the corresponding scaling exponents 1.668, 2.215, 2.166, and 1.351 respectively, which after division by 2 tend eventually to the true Hurst exponent $H = 0.7$ (compare with the previous estimate $\hat{H} = 0.688$ too).

[44] This example shows that, although the variance-time function of the non-stationary Z-process should theoretically be identical to that of the stationary X-process, the actual effect of the harmonic trend is to deform the variance-time plot of simulated *Z*-series, by introducing several power-law scaling regimes which eventually tend to settle on the theoretically anticipated single outer scaling regime. The evolution of scaling regimes, from the inner towards the outer, follows closely the period $T = 50$ of the harmonic trend, with scaling forming in the first half of that period and fading to a “plateau” in the second half. In this example, attainment of the outer scaling regime occurs for aggregation scales greater than about three times the period (i.e. for $m \ge 150$). Similar behavior is found also in variance-time plots obtained from synthetic *Z*-series for several other values of the Hurst exponent (e.g. $H = 0.5$, 0.6, 0.8, 0.9), and of the period of the sinusoidal trend (e.g. $T = 25$, 75, 100), as long as the "signal-to-noise" ratio is near or below unity (i.e. $a / \sigma_X \approx 1$).

[45] When the ratio a / σ_X becomes substantially larger than unity, then the effects of the sinusoidal trend become much stronger, deforming more severely the variance-time plot of the *Z*-process across a much wider range of scales of aggregation. Such an example is shown in Figure 6(f), depicting the variance-time plot of a synthetic series $Z_t = X_t + A_t$, in comparison to that of the *X*-series, where again $X \sim FGN(0, 1, 0.7)$ and A_t has period $T = 50$, but amplitude $a = 10$, so that $a / \sigma_X = 10$. Figure 6(e) depicts the initial section ($1 \le t \le 1000$) of this synthetic time series (of total length $n = 2^{14}$). Comparing the variance-time plots shown in Figures 6(d) and 6(f), one can discern again an inner scaling regime over the same range of small scales ($m < 25$), but no outer scaling regime is immediately evident in 6(f). However, one may argue the existence of an outer scaling regime settling at very large scales of aggregation, by extrapolating the behavior seen in the plotted part of the transient regime ($25 < m < 500$). The argument is as follows. First note that the transient regime consists of a sequence of "funnels" (e.g. $25 < m < 75$, $75 < m < 125$, etc.) forming between consecutive local maxima and an intermediate local minimum. The decreasing branch of each funnel represents a further distortion of a previous plateau (e.g. compare $25 < m < 50$, $75 < m < 100$, etc. in 6(d) and 6(f)), while its increasing branch represents a further distortion of a previous intermediate scaling regime (e.g. compare $50 < m < 75$, $100 < m < 125$, etc. in 6(d) and 6(f)). The sequence of maxima, attained at scales that are integer multiples of the half-period, forms an upper envelope of the transient regime. Evidently, this upper envelope conforms to a power-law with some positive exponent smaller than the scaling exponent $2H$ of the variance-time plot of the *X*-series, and depends on the ratio a / σ_X. At the other end, the sequence of local minima, attained at scales that are integer multiples of the full-period, forms a lower envelope of the transient regime. Therefore, according to an earlier comment, these local minima fall exactly on the variance-time plot of the stationary *X*-series. Consequently, the two envelopes are anticipated to eventually meet and merge after some large multiple of the period (say $m = 1500 = 30$ x 50 in 6(f), instead of $m = 150 = 3$ x 50 in 6(d)). From that scale onward an outer scaling regime will settle in, recovering again the true variance-time plot of both processes *X* and *Z*. This behavior can be verified via synthetic series of much longer length than (say $n=2^{18}$).

[46] In our second example we introduce "intermittence", by modifying the sinusoidal trend A_t so as to become a new deterministic function $B_t = A_t \cdot \sum_{p \in P} I_{[p,p+T/2]}(t)$, where $I_{[p,p+T/2]}(t)$ is the indicator function of the interval [*p, p+T/2*], and *p* assumes values from a finite set of integers $P \subset N$. Specifically, for a synthetic trajectory of length $n = 2^{14}$ (Figure 7a), we consider P = {1, 8000, 16000}, for a given period $T = 50$ (Figure 7b). Non-stationarity is now represented only by very few sinusoidal bursts (e.g. three in the current example), but with much greater ratio $a / \sigma_X = 50$ of amplitude ($a = 50$) versus FGN variance ($\sigma_X = 1$), while almost all the time the new non-stationary process $W_t = X_t + B_t$ coincides with the FGN process *X*. This effect might be reminiscent of the behavior of stormy rainfall with brief intervals of intense activity interrupting long intervals of quiescence (see Figure 7c). Since $W_t^{(m)} = X_t^{(m)} + B_t^{(m)}$, the processes of aggregates $\{W_t^{(m)}\}$ and $\{X_t^{(m)}\}$ have equal variances, and thus identical variance-time function $\sigma^2_{W^{(m)}} = \sigma^2_{X^{(m)}} = m^{2H} \cdot \sigma^2_X$, which is again a power-law function of the scale of aggregation *m*, with exponent $1<2H<2$.

[47] The variance-time plot of a simulated *W*-series, with $X \sim FGN(0, 1, 0.7)$, has been computed for scales of aggregation in the range $1 \le m \le 500$, and it is compared with that of the stationary *X*-process (see Figure 7d). Notably in this case the two variance-time plots do not overlap at all, showing clearly the extremely strong influence of non-stationarity at all scales of aggregation. Two power-law scaling regimes of variance are clearly present again, with scaling exponents estimated by the value 1.875 over the "inner" regime (scales $m < 25$), and by the value 1.015 over the "outer" regime (scales $m > T / 2 = 25$). Note that the exponent of the outer regime yields a very poor estimate ($\hat{H} = 0.507$) of the true Hurst exponent $H=0.7$. We also note that there is no apparent intermediate regime of transience, as the two scaling regimes shape almost perfectly an obtuse angle. This lack of transitional regime matches well the intuitively anticipated abrupt shift, from the strong effect of non-stationary bursts at small scales of aggregation, to the relaxation of such effects at large scales of aggregation, setting by construction the cut-off near the half-period scale $m = T / 2$. Similar behavior is also noted on variance-time plots of synthetic *W*-series obtained for several values of amplitude (e.g. $a = 40$, 75, 100), period (e.g. $T = 25$, 75, 100), and sets of bursts *P*, when superposing a highly intermittent sinusoidal trend to FGN with long memory.

5. Conclusions

[48] The present work aimed primarily to raise some points of skepticism about scientific statements regarding rainfall behavior, when those statements are based merely on mathematical theory, or on analysis of data that may not fulfill fundamental theorized assumptions, e.g. stationarity. We have shown that the theorized assumption of stationarity must be tested thoroughly and much more rigorously from a statistical standpoint. Formal statistical procedures, suitable for testing homogeneity of mean, variance, and covariance, were applied to Maghera and Firenze data, raising reasonable doubts about the extent up to which stationarity is a safe and valid assumption to adopt with regard to those data. Specifically, homogeneity of mean holds within individual monthly sections (Table 1), but not inter-annually for any given month of the year (Table 1 and Figures 1-2). Furthermore, homogeneity of variance and covariance is rejectable with at least 95% confidence over monthly, annual, and multi-yearly (i.e. global) time scales, for the majority of sections representing those time scales in the data (Tables 2-5), while scale-dependence of stationarity is evident over shorter than monthly time scales (Figures 4-5).

[49] Subsequently, we concerned ourselves with the effects that certain types of deterministic trends bear on the variance-time plot, when superposed on stationary processes whose variance-time plot is a single power-law function of scale of aggregation (e.g. on FGN processes with long memory). These examples demonstrate very clearly that variance-time plots obtained from raw data, without previous exploration of possible sources of non-stationarity present in them, can be extremely misleading if used as a tool for identifying regimes of scales of aggregation over which the variance of aggregates may be scaled by power-law scaling. Specifically, estimated variance-time plots exhibit several power-law scaling regimes, separated by intermediate transitional regimes, while their theoretical counterparts are indeed single power-law functions of the scale of aggregation (recall that $\sigma^2_{W^{(m)}} = \sigma^2_{Z^{(m)}} = \sigma^2_{X^{(m)}} = m^{2H} \cdot \sigma^2_X$). The sharpness or subtlety of such distortions is appreciably controlled, by tuning certain features (e.g. signal to noise ratio and intermittence) built into these examples. In any event, these distortions are plainly artifacts of the deterministic harmonic trend, and they do not characterize the stochastic behavior of the underlying aggregated non-stationary processes (i.e. *Z* and *W*). The examples considered above, bear no direct connection with rainfall dynamics, and they are not portrayed here as models of rainfall at any scale,

although they were motivated by an interest to imitate apparent scale-dependence of stationarity in rainfall time series, and also to demonstrate variance-time plots (e.g. Figure 7(d)) similar to those obtained by *Marani* [2003] from rainfall time series. In light of these examples, one may safely conclude that even if the regimes detected by *Marani* [2003] are not artifacts of intermittent non-stationarities, they are not signaling an exclusively special property of temporal rainfall processes.

Acknowledgments

The authors would like to thank Ente Zona Industriale di Porto Marghera (Via delle Industrie, 19 - 30175 Porto Marghera, Venezia, Italy) for providing the rainfall data of Marghera gaging station. The support of the National Research Council of Italy, through CNR-GNDCI contract n.03.00065.GN42 is acknowledged.

References

1. Ahamada I., and M. Boutahar, (2002) Tests for covariance stationarity and white noise, with an application to euro/us dollar exchange rate. An approach based on the evolutionary spectral density, *Economics Letters*, 77, 177–186.

2. Beran J. (1994*) Statistics for long-memory processes*, Chapman & Hall.

3. Bhattacharya R.N. and Waymire E.C. (1990) *Stochastic processes with applications*, J. Wiley & Sons.

4. Bhattacharya R.N., Gupta V.K., and Waymire E.C. (1983) The Hurst effect under trends, *J. Appl. Probab.,* 20, 649-662.

5. Bras R. L., and I. Rodriguez-Iturbe, (1985) *Random function and hydrology*, Addison-Wesley.

6. Brockwell P.J. and Davis R.A. (1991) *Time Series: Theory and Methods, 2nd Edition*, Springer-Verlag.

7. Dang T. D., and S. Molnár, (1999) On the effects of non-stationarity in long-range dependence tests, *Periodica Polytechnica* SER. EL. ENG. 43(4) 227-250.

8. Davis A., A. Marshak, W. Wiscombe, and R. Cahalan, (1996) Scale invariance of liquid water distribution in marine stratocumulus, I, Spectral properties and stationarity issues, *J. Atmos. Sci.*, 53, 1538– 1558.

9. Diebold F.X. and Inoue A. (2001) Long memory and regime structuring, *J. Econometrics*, 105, 131-159.

10. Embrechts P. and Maejima M. (2002) *Selfsimilar processes*, Princeton Univ. Press.

11. Fabry F., (1996) On the determination of scale ranges for precipitation fields, *J. Geophys. Res.* 101 12819–12826.

12. Fraedrich K, and C. Larnder, (1993) Scaling regimes of rainfall time-series, *Tellus* 45A, 289–298.

13. Georgakakos K.P., Carsteanu A.A., Sturdevant P.L., and J.A. Cramer, (1994) Observation and analysis of Midwestern rain rates. *J. App. Met.* 33 1433–1444.

14. Gilbert R.O. (1987) *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand Reinhold.

15. Harris D., Menabde M., Seed A., and G. Austin, (1996) Multifractal characterization of rain fields with a strong orographic influence, *J. Geophys. Res.* 101 26405–26414.

16. Kärner O., (2002) On nonstationarity and antipersistency in global temperature series, *J. Geophys. Res.*, 107 (D20), 4415, doi:10.1029/2001JD002024,2002.

17. Kottegoda N. T., (1985) Assessment of non-stationarity in annual seris through evolutionary spectrum, *J. Hydrol.*, 76, 381-402.

18. Künsch H. (1986) Discrimination between monotonic trends and long-range dependence, *J. Appl. Probab.,* 23, 1025-1030.

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

20. Lo A.W. (1987) Long-term memory in stock market prices, Unpublished Manuscript, Sloan School of Management, MIT, Cambridge, Massachusetts.

21. Malamud B.D. and D. L. Turcotte, (1999) Self-affine time series: measures of weak and strong persistence, *J. Stat. Plan. Inf.* (80) 173-196.

22. Mandelbrot B. B. (1963) The variation of certain speculative prices, *Journal of Business*, 36, 394-419.

23. Mandelbrot B. B., (1983) *The Fractal Geometry of Nature*, W.H. Freeman, New York.

24. Mandelbrot B. B., (1998) Multifractals and 1/f noise, Springler-Verlag, New York.

25. Marani M., (2003) On the correlation structure of continuous and discrete point rainfall, *Wat. Resour. Res.*, 39(5), 1128, doi:10.1029/2002WR001456,2003.

26. Mikosch T. and Stărića C. (2000) Is it really long memory we see in financial returns? In, Embrechts P. (Ed.) *Extremes and integrated risk management*, Risk Books, Risk Waters Group, pp. 149-168.

27. Mikosch T. and Stărića C. (2001) Long range dependence effects and ARCH modelling. In, Doukhan P., Oppenheim G., and Taqqu M.S. (Eds.) *Long range dependence*, Birkhäuser, Basel.

28. Olsson J., and P. Burlando, (2002) Reproduction of temporal scaling by a rectangular pulses rainfall model, *Hydrol. Process.*, 16, 611–630.

29. Pagan A., and G. Schwert, (1990) Testing for covariance stationarity in stoch market data, *Economics Letters*, 33, 165–170.

30. Paxson V., (1995) Fast approximation of self similar network traffic*, Tech. Rep. LBL-36750*, 1995.

31. Phillips P.C.B. (1987) Time series regression with a unit root, *Econometrica*, 55, 277-301.

32. Priestley M.B., (1965) Evolutionary spectra and non-stationary processes, *J. R. Statist. Soc.*, *Series B*, 27, 204-237.

33. Priestley M.B., (1981-a) *Spectral Analysis and Time Series, Volume 1*, Academic Press.

34. Priestley M.B., (1981-b) *Spectral Analysis and Time Series, Volume 2*, Academic Press.

35. Priestley M.B., and T. Subba Rao, (1969) A test for non-stationarity of time series, *J. R. Statist. Soc.*, 31, 140-149.

36. Samorodnitsky G. and Taqqu M.S. (1994) *Stable non-Gaussian random processes*, Chapman & Hall.

37. Tessier Y., Lovejoy S., Hubert P., Schertzer D., and S. Pecknold, (1996). Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions, *J. Geophys. Res.* 101 26427–26440.

38. Teverovsky V., and M.S. Taqqu, (1997) Testing for long range dependence in the presence of shifting means or a slowly declining trend, using a variance-time estimator, *Time Series Analysis*, 18(3) 279-304.

39. Yaglom A.M. (1987) *Correlation Theory of Stationary and Related Random Functions I: Basic Results*, Springer-Verlag.

40. Whitcher, B., S. D. Byers, P. Guttorp, and D. B. Percival (2002). Testing for Homogeneity of Variance in Time Series: Long Memory, Wavelets and the Nile River. *Wat. Resour. Res.*, 38 (5), 10.1029/2001WR000509.

41. Wilby R.L. (1997) Non-stationarity in daily precipitation series: Implications for GCM down-scaling using atmospheric circulation indices, *International Journal of Climatology*, 17, 439-454.

LIST OF TABLES

LIST OF FIGURES

	January	February	March	April	May	June	July	August	September	October	November	December
1993	0.001	0.013	0.049	0.062	0.016	0.050	0.154	0.027	0.148	0.109	0.101	0.038
1994	0.102	0.058	0.001	0.162	0.060	0.038	0.103	0.074	0.153	0.087	0.088	0.019
1995	0.053	0.111	0.084	0.089	0.223	0.276	0.075	0.121	0.165	0.030	0.010	0.189
1996	0.094	0.072	0.016	0.250	0.080	0.098	0.044	0.126	0.096	0.188	0.079	0.228
1997	0.109	0.007	0.016	0.068	0.050	0.084	0.133	0.045	0.041	0.057	0.183	0.119
1998	0.043	0.011	0.015	0.159	0.079	0.137	0.062	0.024	0.162	0.229	0.022	0.020
1999	0.045	0.028	0.056	0.126	0.044	0.191	0.137	0.039	0.044	0.157	0.218	0.091
2000	0.001	0.010	0.075	0.077	0.087	0.037	0.056	0.107	0.111	0.216	0.190	0.073
$\overline{\mu}_X$	0.056	0.039	0.039	0.124	0.080	0.114	0.096	0.070	0.115	0.134	0.111	0.097
$s(\overline{\mu}_X)$	0.043	0.038	0.031	0.065	0.062	0.084	0.042	0.042	0.051	0.075	0.078	0.078
$\overline{\mu}_X \pm z_{0.975} \cdot s(\overline{\mu}_X)$	[0.026, 0.086]	[0.013, 0.065]	[0.017, 0.060]	[0.079, 0.169]	[0.037, 0.123]	[0.056, 0.172]	[0.067, 0.125]	[0.041, 0.100]	[0.080, 0.150]	[0.082, 0.186]	[0.057, 0.165]	[0.043, 0.151]

Table 1. Slopes of mean-time plots for every month and year, from 1993 to 2000, Marghera. Moreover, for each month of the year, the sample mean of slopes across years is reported, along with its standard error and 95% confidence interval.

	January	February	March	April	May	June	July	August	September	October	November	December	Year
1993	3.2	0.0	1.1	3.6	0.0	22.3	4.6	0.2	0.4	8.6	0.3	0.5	9.7
1994	0.0	5.1	0.0	0.1	32.7	17.0	12.0	0.0	24.0	0.0	1.7	20.3	2.5
1995	15.3	0.0	0.6	0.4	0.0	11.0	0.2	0.2	10.4	0.0	0.4	0.0	1.1
1996	10.6	0.0	27.0	0.0	0.2	1.5	3.0	8.5	0.0	0.0	2.0	14.5	10.0
1997	0.0	30.0	0.0	0.0	20.5	16.5	32.0	17.0	10.0	17.5	0.2	0.4	2.1
1998	11.0	3.0	0.0	5.0	0.0	0.0	23.0	0.0	0.3	0.0	0.1	22.3	1.1
1999	8.0	9.5	5.7	3.0	27.5	21.0	17.0	14.0	4.8	6.0	1.2	0.0	1.2
2000	0.5	24.5	0.0	0.0	5.7	28.0	13.5	2.8	0.0	0.2	1.7	0.0	7.8
All years	**2.8**												

Table 2. p-values of CUSUM stationarity tests for every month and year, from 1993 to 2000, along with p-value for each annual series (last column), and p-value over the entire record (last row), Marghera.

	January	February	March	April	May	June	July	August	September	October	November	December	Year
1962	13.8	0.0	3.0	2.4	5.7	0.5	1.0	NaN	8.0	0.0	0.3	0.0	29.7
1963	0.0	7.1	0.0	5.0	7.5	6.5	24.5	14.7	2.0	0.0	2.7	0.0	0.5
1964	1.9	0.5	0.0	0.0	0.0	15.2	11.4	12.6	3.6	0.0	0.0	0.0	2.7
1965	0.0	0.0	0.0	0.2	0.0	0.0	1.3	0.7	8.6	0.0	5.3	0.5	0.0
1966	1.5	0.1	10.0	27.0	4.5	23.5	10.2	0.0	0.0	0.1	0.1	0.0	0.1
1967	0.0	0.5	0.2	6.0	4.8	18.4	15.2	2.0	0.1	0.0	0.0	0.0	4.9
1968	0.0	1.0	3.9	0.0	0.0	3.3	1.5	0.0	12.5	6.3	0.1	0.1	0.2
1969	1.8	3.0	1.0	0.0	5.8	0.0	10.0	8.0	13.6	0.0	25.0	0.3	1.3
1970	29.7	1.0	0.0	0.2	15.0	0.7	0.1	1.5	7.5	0.0	0.4	0.0	1.1
1971	0.1	0.0	2.5	4.5	0.0	0.5	17.0	13.5	15.2	1.5	1.4	0.0	8.6
1972	0.0	2.5	0.8	0.0	4.9	0.2	18.0	0.1	0.0	0.3	16.8	0.0	0.0
1973	0.1	0.0	0.1	0.2	NaN	NaN	NaN	NaN	0.1	7.4	0.1	0.0	0.0
1974	0.0	0.0	0.0	0.0	0.4	4.0	2.5	0.4	0.0	12.0	3.5	0.0	0.9
1975	0.0	0.0	4.5	0.0	19.0	0.0	0.0	9.8	0.1	0.1	0.4	0.0	1.1
1976	0.0	0.0	0.0	0.0	0.0	0.0	24.5	18.4	0.5	0.0	0.1	0.0	0.1
1977	0.1	0.7	0.0	0.1	19.5	0.1	0.0	16.8	0.0	0.0	18.4	0.0	10.4
1978	0.0	0.0	3.0	8.0	0.0	4.5	0.0	0.9	0.0	0.0	0.0	9.8	0.0
1979	0.0	0.0	1.5	0.3	NaN	0.2	0.0	12.0	0.4	23.0	0.2	0.1	10.4
1980	0.4	0.0	3.5	0.2	0.0	0.1	13.0	13.5	2.1	10.4	0.0	0.0	5.9
1981	0.0	0.0	0.0	0.0	2.0	1.5	20.0	25.0	8.0	0.0	0.0	0.4	0.4
1982	0.0	0.0	6.5	4.0	6.0	1.5	0.1	25.5	0.0	0.3	2.5	2.1	0.0
1983	18.3	0.0	0.0	0.0	0.1	0.0	1.2	0.8	0.0	0.1	0.0	0.0	16.8
1984	0.1	0.0	0.0	0.0	0.2	0.0	27.0	5.4	3.0	0.0	1.1	3.5	0.8
1985	0.1	3.8	0.0	4.0	0.0	0.1	0.0	4.0	29.7	0.0	0.0	0.0	15.2
1986	0.0	0.1	0.0	2.5	1.1	0.0	0.2	0.0	2.3	0.0	0.0	0.0	7.8
All years	**1.0**												

Table 3. p-values of CUSUM stationarity tests for every month and year, from 1962 to 1986, along with p-value for each annual series (last column), and p-value over the entire record (last row), Firenze.

	January	February	March	April	May	June	July	August	September	October	November	December	Year
1993	2.0	19.9	16.6	1.5	21.7	88.0	17.2	0.0	0.6	25.9	0.1	49.2	0.0
1994	0.0	15.1	5.4	0.1	13.3	87.1	31.7	0.4	23.3	3.8	33.9	20.5	1.5
1995	0.0	1.3	2.8	0.2	2.2	3.8	6.3	40.5	50.1	66.9	18.4	0.1	11.9
1996	27.1	15.3	4.3	1.9	54.9	0.2	15.6	60.9	31.4	15.6	32.1	16.9	0.0
1997	0.2	71.7	0.7	0.8	18.2	37.8	47.6	71.0	14.3	26.4	18.6	54.0	1.7
1998	3.6	11.5	36.4	12.5	13.7	3.5	20.6	2.4	12.8	0.1	1.9	14.6	2.2
1999	42.7	18.6	11.3	4.6	8.5	12.3	18.1	45.9	0.1	0.0	28.6	8.9	22.1
2000	0.2	0.5	0.3	0.0	52.1	9.2	34.9	8.1	15.1	0.0	15.6	0.3	0.0
All years	**1.1**												

Table 4. *p*-values of covariance stationarity tests (Ahamada & Boutahar test) for every month and year, from 1993 to 2000, along with *p*-value for each annual series (last column), and *p*-value over the entire record (last row), Marghera.

	January	February	March	April	May	June	July	August	September	October	November	December	Year
1962	0.0	0.0	0.0	0.0	6.8	0.1	10.7	NaN	0.0	0.0	0.0	0.0	0.0
1963	0.0	0.0	0.0	0.0	3.1	0.0	4.2	31.0	0.0	0.0	0.0	1.8	0.0
1964	0.0	0.0	0.0	0.0	0.1	0.2	14.1	3.5	0.0	0.6	0.0	0.0	0.0
1965	0.3	0.1	0.0	0.0	0.0	0.0	60.1	0.0	0.0	1.0	0.0	0.0	0.0
1966	0.4	0.0	0.0	0.3	0.1	3.1	0.1	0.0	0.0	0.0	0.3	0.0	0.0
1967	0.0	0.0	0.2	15.7	0.0	0.4	42.7	5.5	0.5	0.0	0.0	0.0	0.0
1968	0.0	0.2	1.4	0.2	0.0	3.4	0.0	8.7	33.1	0.0	0.0	0.0	0.0
1969	0.0	0.0	0.8	0.3	0.3	0.0	2.5	0.0	0.0	10.5	0.0	0.1	0.0
1970	1.3	0.0	0.0	1.0	5.3	3.8	43.9	1.4	22.4	15.0	0.0	0.0	0.0
1971	0.0	0.7	0.0	3.5	0.0	0.1	19.9	2.2	0.7	1.9	0.0	0.0	0.0
1972	0.0	9.8	0.0	0.0	0.0	0.0	1.0	0.0	0.0	3.5	0.0	0.1	0.0
1973	0.0	0.0	0.0	0.0	NaN	NaN	NaN	NaN	0.0	0.0	0.0	0.7	0.0
1974	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	3.7	0.2	1.1	0.0
1975	0.0	0.0	0.1	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
1976	1.0	0.0	0.5	0.0	0.1	0.0	10.1	0.7	0.0	4.4	0.0	0.0	0.0
1977	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.2	0.0	0.1	0.0
1978	0.6	0.0	0.0	0.2	0.3	0.0	0.0	1.3	18.6	0.0	0.0	0.0	0.0
1979	0.0	0.0	0.0	0.0	NaN	0.1	0.0	0.5	0.0	0.0	0.0	0.0	0.0
1980	2.1	0.0	0.0	0.0	5.2	0.0	3.6	0.4	39.7	0.1	0.0	0.0	0.0
1981	0.0	0.0	0.0	0.0	1.1	0.0	2.1	14.3	0.0	0.0	1.4	0.3	0.0
1982	0.2	0.0	4.3	0.4	0.0	0.0	0.0	0.1	5.7	0.1	0.0	0.0	0.0
1983	51.6	0.0	0.0	0.0	0.0	0.0	0.3	0.0	16.1	0.0	0.0	0.0	0.0
1984	0.0	0.0	0.0	0.0	0.1	0.0	6.0	0.2	0.0	0.0	0.0	2.5	0.0
1985	0.0	0.7	0.0	0.0	0.0	0.0	4.6	20.0	67.4	0.0	0.0	0.0	0.0
1986	0.0	0.0	0.0	0.0	2.7	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0
All years	**0.0**												

Table 5. *p*-values of covariance stationarity tests (Ahamada & Boutahar test) for every month and year, from 1962 to 1986, along with *p*-value for each annual series (last column), and *p*-value over the entire record (last row), Firenze.

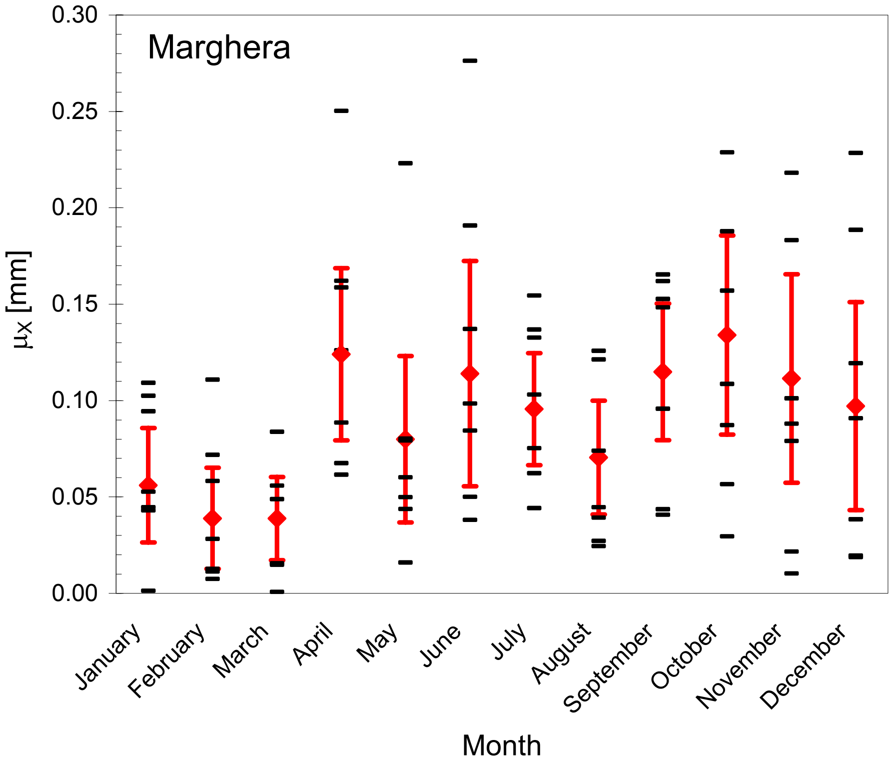


Figure 1. 95% confidence interval of the mean of hourly rainfall, μ_X, for each month and comparison with sample estimates, Marghera.

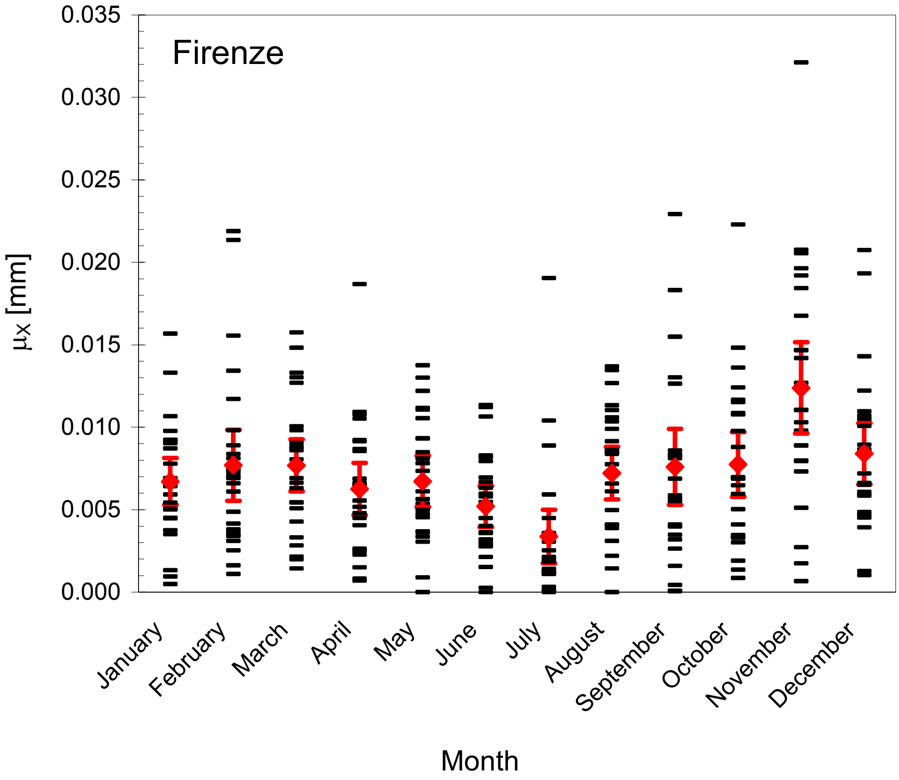


Figure 2. 95% confidence interval of the mean of 5 minutes rainfall, μ_X, for each month and comparison with sample estimates, Firenze.

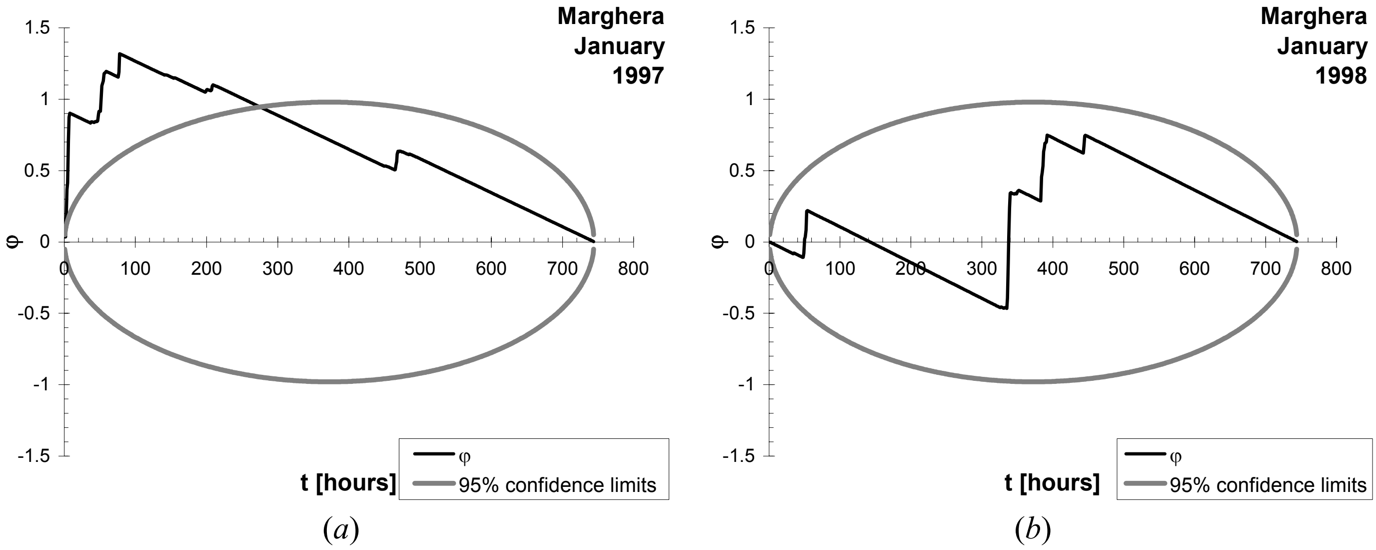


Figure 3. CUSUM stationary test for January hourly rainfall at Marghera: (*a*) for 1997 year and (*b*) for 1998 year. The plot shows the sampling Pagan & Schwert [1990] statistics versus time, and its associated rejection region (5% level of significance).

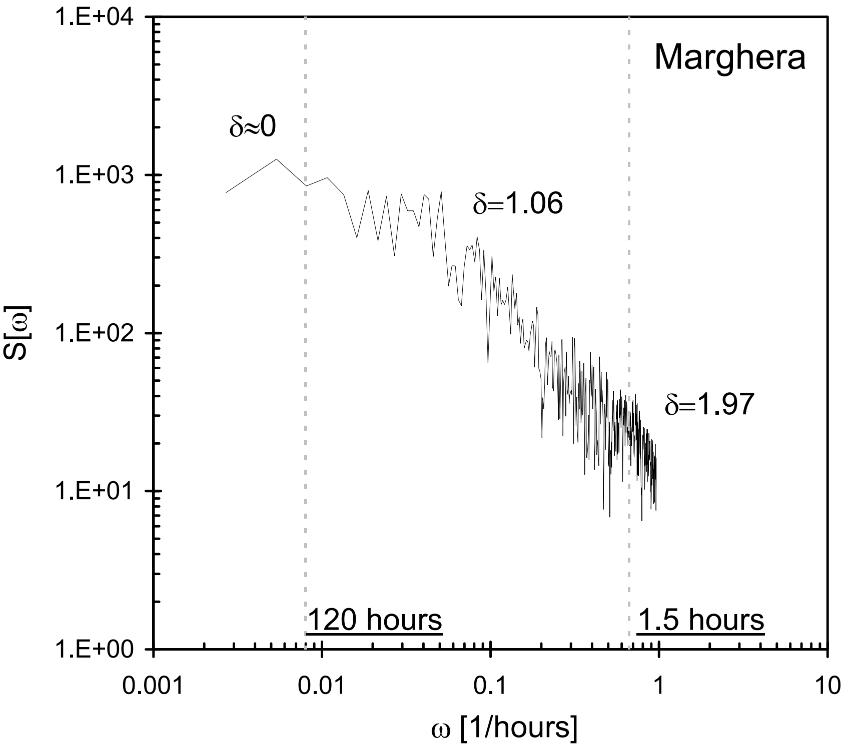


Figure 4. Average power spectrum of January hourly rainfall data, Marghera.

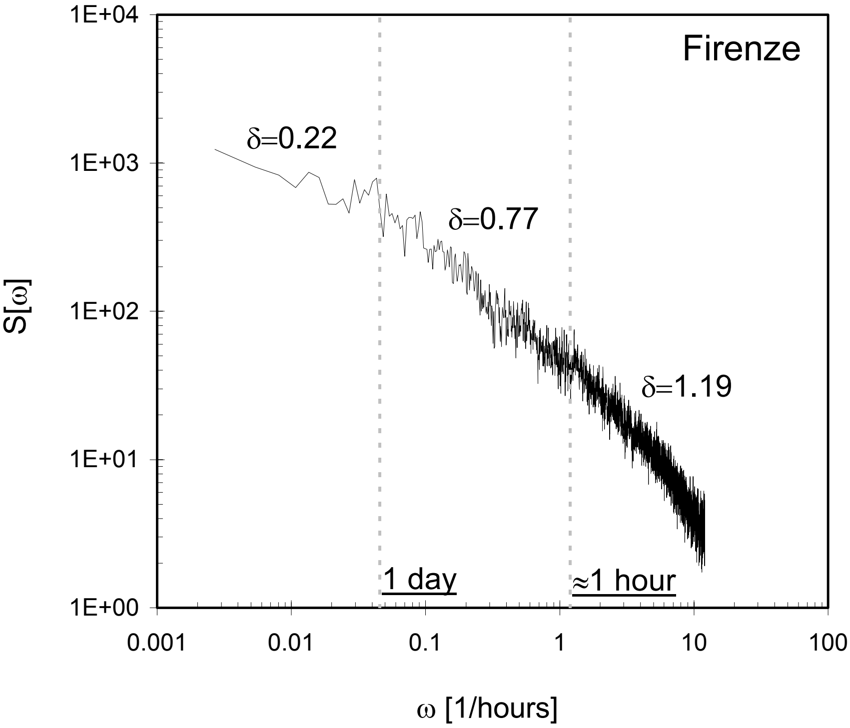


Figure 5. Average power spectrum of January 5-minutes rainfall data, Firenze.

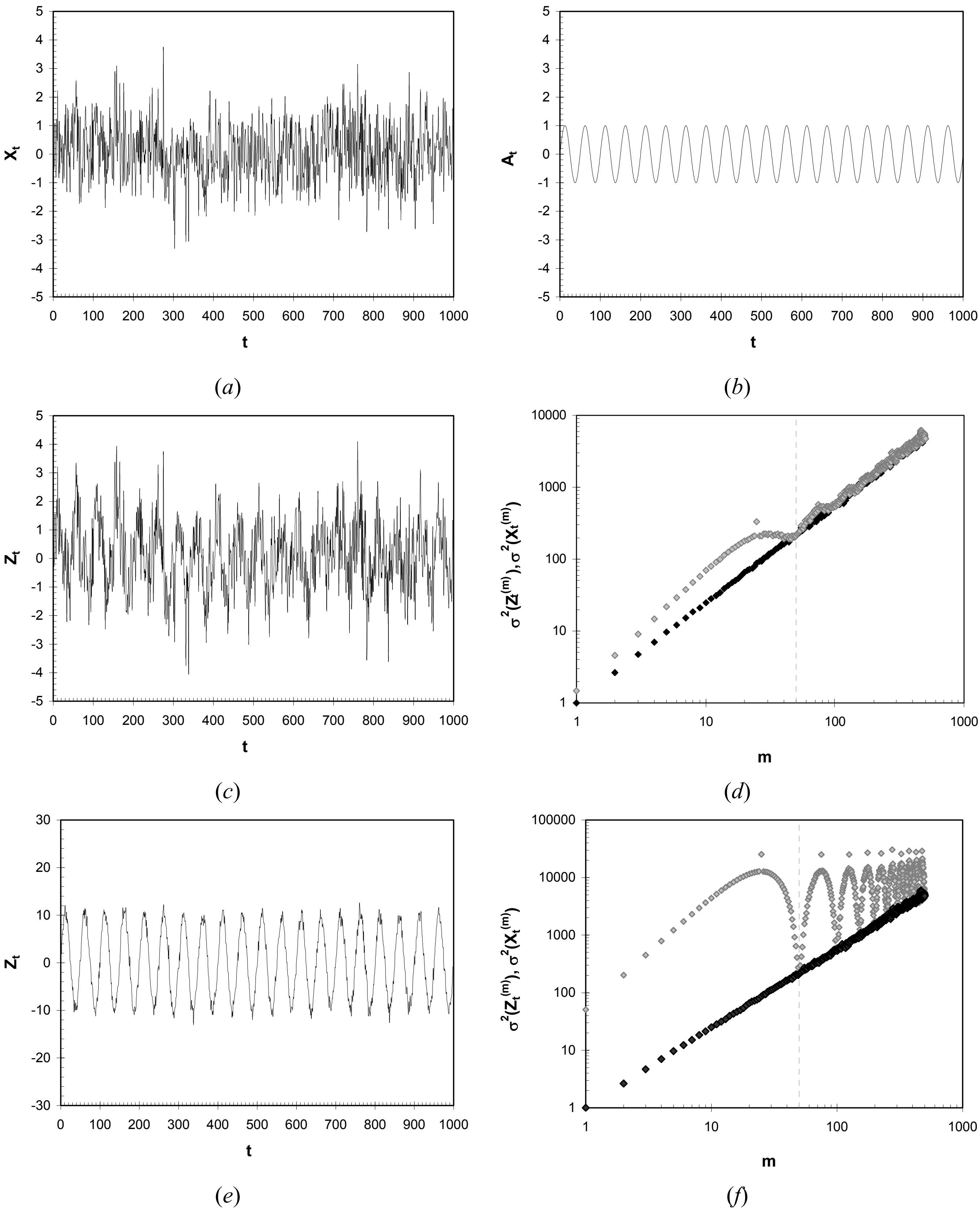


Figure 6. Datagrams of synthetic data, and variance-time plots: (*a*) $X_t \sim$ FGN(0, 1, 0.7). (*b*) $A_t = \sin(2\pi t/50)$. (*c*) $Z_t = X_t + A_t$. (*d*) Variance-time plot of X_t (black dots), and of $Z_t = X_t + A_t$ (gray dots). (*e*) $Z_t = X_t + 10A_t$. (*f*) Variance-time plot of X_t (black dots), and of $Z_t = X_t + 10A_t$ (gray dots).

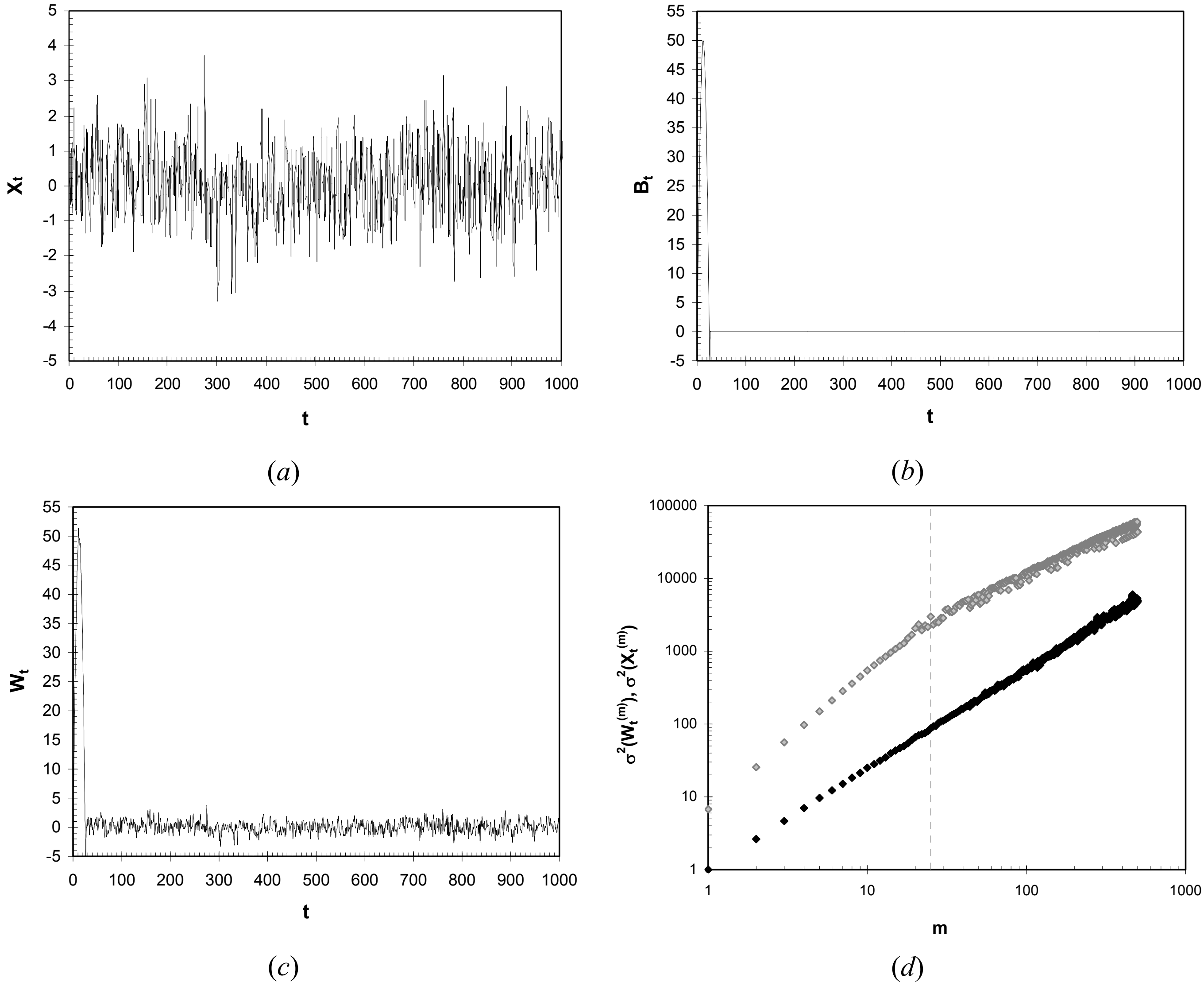


Figure 7. Datagrams of synthetic data and variance-time plots: (*a*) X_t ~ FGN(0, 1, 0.7). (*b*) B_t is an "intermittent" sinusoidal function. (*c*) $W_t = X_t + B_t$ (superposition of the "intermittent" sinusoidal trend to the Fractional Gaussian Noise). (*d*) Variance-time plot of X_t (black dots), and of W_t (gray dots).