Climate change, Hurst phenomenon, and hydrologic statistics

Demetris Koutsoyiannis

Department of Water Resources, Faculty of Civil Engineering,

National Technical University, Athens

Heroon Polytechneiou 5, GR-157 80 Zographou, Greece

(dk@hydro.ntua.gr)

Paper submitted to Water Resources Research

May 2001

Abstract. The intensive research of the recent years on climate change may have resulted in diverging model predictions of the future climate and, also, have caused scientific debate on the detection and attribution of climate changes. Undoubtedly, however, it has leaded to the strong conclusion that climate has ever, through the planet history, changed irregularly on all time scales. The changes of the climate on all scales are closely related to the Hurst phenomenon, which has been detected in almost every long hydroclimatic time series and is nothing more than a simple scaling behavior of climate variability over timescale. The climate variability, anthropogenic or natural, increases the uncertainty of the hydrologic processes. It is shown that hydrologic statistics, the branch of hydrology that deals with uncertainty, in its current status is not consistent with the varying character of climate and more specifically with the Hurst phenomenon. Typical statistics used in hydrology such as means, variances, cross- and auto-correlations and Hurst coefficients, and the variability thereof, are revisited under the hypothesis that climate changes on all scales, following a simple scaling law, and new estimators are developed which in many cases differ dramatically from the classic statistical estimators. The new statistical framework is applied to real-world examples for typical tasks such as estimation and hypothesis testing where again the results depart significantly from those of the classic statistics.

GAP index terms: 1869 Stochastic processes, 1833 Hydroclimatology, 1620 Climate dynamics, 6309 Decision making under uncertainty

1. Introduction

In the last two decades, climate change has been the subject of intensive scientific research, focusing on the understanding of factors, mechanisms and processes related to climate, on modeling the climate at the global scale using the so called climatic, or general circulation models, and on the detection and attribution of changes in the past climate. Climate change has been, at the same time, the subject of scientific debate concentrating firstly on whether existing climatic records indicate a significant change of the present climate versus the past, and, secondly, on whether detected changes are attributed to natural or anthropogenic forcings.

Thus, there is a number of studies speaking about "unprecedented global warming" of the past two decades, which must be attributed to anthropogenic forcings, such as the emissions of CO_2. To refer to a recent example, *Stott et al.* [2000] comparing observations with simulations of a coupled ocean-atmosphere general circulation model conclude that both natural and anthropogenic factors have contributed significantly to 20th century temperature changes. More specifically, when they employed only natural forcings in their model, predictions did not match up well with the observational temperature record in the last 35 years, where the calculated temperatures fell somewhat below the measured temperatures (although the model results matched up observations for the earlier years starting from 1860). Finally, they predict that anthropogenic global warming under a standard emissions scenario will continue at a rate similar to that observed in recent decades.

On the other side, to invoke another recent study, *Przybylak* [2000] detects no global warming in the recent years. Specifically, he studies mean monthly temperatures of Arctic and sub-Arctic areas using observations and grid data over the period of instrumental observations. According to the author, the polar regions “should play a very important role in the detection of global changes” as they are the most sensitive and “warming and cooling epochs should be seen most clearly here and should also occur earlier than in other parts of the world." The analyses show that in the Arctic, "since the mid-1970s, the annual temperature shows no clear trend" whereas "the highest temperatures since the beginning of instrumental observation occurred clearly in the 1930s" and "even in the 1950s the temperature was higher than in the last 10 years"; in addition, "the level of temperature in Greenland in the last 10-20 years is similar to that observed in the 19th century". The author concludes that "the observed variations in air temperature in the real Arctic are in many aspects not consistent with the projected climatic changes computed by climatic models for the enhanced greenhouse effect," and "the temperature predictions produced by numerical climate models significantly differ from those actually observed."

The issue of climate modeling capability is also examined in the comprehensive study by *Barnett et al.* [1999] who state that, at present, it is not possible to distinguish the relative contributions of specific natural and anthropogenic forcings to observed climate change. One of the main reasons is that fully realistic simulations of climate change due to the combined effects of all anthropogenic and natural forcings mechanisms have not been computed yet. The authors also state that "there has been to date no completely convincing demonstration that the anthropogenic effects predicted by advanced climate models have been unambiguously detected in observations", whereas "given the large model uncertainties and limited data, a reliable weighting of the different factors contributing to the observed climate change cannot currently be given" and "we cannot attribute, at this time, with a high level of statistical significance, the observed changes in global and large-scale regional climate to anthropogenic forcing alone." They conclude that "the current state of affairs is not satisfactory."

If the current state of affairs is not satisfactory when dealing with the past and present climate, things are even worse when attempting to produce predictions of the future. (After all, it is not so difficult to fit models with a number of adjustable parameters to historical records, but in predicting the unknown future the situation is completely different.) Climatic models are necessarily simplified representations of the climate system as they do not describe completely the dynamics of all involved processes but, rather, they use simplified representations, known as parameterizations for a large number of processes including mixing, convection and clouds [e.g. *von Storch et al.*, 2001]. On the other hand, climate is influenced by several factors with opposing effects [e.g., *Ledley et al.*, 1999], either anthropogenic (greenhouse gases, aerosols) or natural (solar irradiance, clouds, hydrological cycle, volcanic activity, etc.). Some of these factors are extremely difficult or even impossible to incorporate in models and/or to predict. Overall, as *von Storch et al.* [2001] put is, "climate must be considered as a stochastic system, and our climate simulation models as random number generators". Even merely the natural forcings are enough to result in a perpetually changing climate. Indeed, we now know that climate "changes irregularly, for unknown reasons, on all timescales" [*National Research Council*, 1991, p. 21].

The uncertainty or unpredictability becomes even higher when moving from general climatic variables, such as temperature (which is the key variable for most of the studies mentioned above) to hydrologic variables such as rainfall and runoff, and from the coarse spatial scale of climatic models to the finer spatial scale of hydrologic models. In parallel, the importance of these hydrologic variables is higher, when dealing with engineering and management issues such as design and operation of projects and hydrosystems.

In the last two decades, hydrologists have developed different strategies or methodologies to deal with climate change. One of these consists of the so-called downscaling of climatic models results into the area and timescale of interest and, subsequently, the feeding of hydrologic models with the downscaled data. Another option is the adoption of alternative scenarios with plausible shifts in the average hydrologic regime of the area of interest (often expressed as a percentage of the historical averages). The first method suffers from the accumulation of uncertainties and errors in the chain of models that are used, whereas the second one is arbitrary and lacks quantification of uncertainty.

Traditionally, however, hydrologic and water resources studies have been based on the quantification of the natural uncertainties and the resulting risk in terms of probability. The discipline of hydrologic statistics [e.g. *Chow*, 1964; *Yevjevich*, 1972; *Haan*, 1977; *Kottegoda*, 1980; *Kite*, 1988; *Hirsch et al.*, 1993; *Stedinger et al.*, 1993] has been well developed and applied in almost every hydrologic engineering study. Hydrologic statistics, though, has been based on the implicit assumption of a stable climate. The questions arise then: (1) Is hydrologic statistics, in its present state, consistent with the assumption of a varying climate? (2) If not, what adaptations are needed to achieve this consistency? (3) Can hydrologic statistics be used to quantify the total uncertainty under a varying climate?

These are the main questions studied in this paper. The usefulness of the answers to these questions from an engineering and management point of view is almost obvious. However, one may argue that the anthropogenic climate change cannot be predicted at all with statistical means, based on historical records that are almost free of anthropogenic influences. This may be correct if indeed, the contribution of anthropogenic forcings is high, relative to that of the natural forcings, a statement that is still unproven. Even if this is the case, the usefulness of statistical estimators consistent with a naturally varying climate is high for the procedure of detecting climate change. Detection of climate change requires demonstrating that the observed change is larger than would be expected to occur by natural causes alone, and this is clearly a statistical problem. In addition, even when working with purely deterministic climatic models, the usefulness of statistics is undeniable in the phase of evaluating the model results.

If we revisit the classic statistical estimators (e.g. for the mean, variance, etc.) and glance at their mathematical proofs, we will see that they are based on the assumption that statistical samples consist of independent, identically distributed variables. Obviously, this assumption is not consistent with the nature of hydrometeorological time series. What is a consistent assumption is studied in section 2 by refining older ideas on the stochastic representation of hydrometeorological processes and with the help of real world examples. The finally adopted stochastic representation is consistent with the aspect of a varying climate. This stochastic representation influences seriously typical statistical tasks such as estimation, prediction and hypothesis testing. This is studied in detail in section 3 where the inappropriateness of classic estimators is demonstrated and new statistical estimators and confidence limits of estimations are derived. The developed statistical framework is further demonstrated in section 4 by means of case studies. The conclusions are drawn in section 5.

2. Stochastic representation of hydrometeorological processes

2.1 Older ideas in light of a varying climate

Hydrometeorological processes such as rainfall, runoff, evaporation, etc., have been often considered [e.g. *Haan*, 1977, p. 275; *Kottegoda*, 1980, p. 26, *Salas*, 1993, p. 19.7; *Shaw*, 1994, p. 372] as composing of at least three parts, a deterministic part (T) representing trends and jumps, another deterministic part representing periodicity (P), and a stationary stochastic part (Ξ). Moreover, the process of interest, say X_i with i denoting discrete time, has been decomposed [*Kottegoda*, 1980, p. 26, *Shaw*, 1994, p. 373] according to

$$X_i = T_i + P_i + \Xi_i \tag{1}$$

This conceptualization originates from observation of real-world time series, which indeed exhibit parts with rising or falling trends; the periodicity is obvious due to the annual cycle of the earth movement; and the stochastic part is the remaining part that is unexplained in a deterministic framework.

We maintain that this conceptualization is inappropriate and inconsistent for hydrometeorological processes. To demonstrate this, let us start with the periodicity term, assuming for the time being that $T_i = 0$. If P_i is a purely deterministic periodic component, it becomes clear from (1) that $X_i - P_i \equiv \Xi_i$ will have, for instance, constant variance, independent of the season i (e.g., month), which is not correct: it is well known that the variance of hydrometeorological processes depends on the season. A better alternative is to consider $X_i = P_i + a_i \Xi_i$ also assuming that a_i is a deterministic periodic term. If for example we assume that P_i equals the seasonally varying mean, a_t equals the seasonally varying standard deviation, and let Ξ_i have zero mean and unit standard deviation, then the latter model can represent well the seasonally varying mean and standard deviation of X_i. However, it can be easily shown that it cannot represent other important seasonally varying statistics, such as coefficients of autocorrelation and skewness, which the model treats as being constant for all seasons. Therefore, a decomposition of the original process X_i into a periodic deterministic and a nonperiodic stochastic part is not possible. This, however, is not a major problem: The theory of stochastic process can very well handle stochastic processes with intrinsic periodicity, using the so called cyclostationary (also called periodic or seasonal) modeling tools [see e.g. *Gardner*, 1989, pp. 323-404; *Bras and Rodriguez-Iturbe*, 1985; *Salas*, 1993] or using disaggregation techniques [e.g., *Koutsoyiannis*, 2001a]. Therefore, we will not examine further in this paper the periodicity issue, which has been effectively resolved, but not in terms of decomposition or isolation of the periodic part.

Now, let us come to the trend term, which is the most significant for the scope of this paper. We use three example time series, which are depicted in Figure 1 through Figure 3. The first example, shown in Figure 1, is a long time series (992 years) representing the North Hemisphere temperature anomalies (in ^{o}C versus 1961-90 mean). This series (available from ftp.ngdc.noaa.gov/paleo/contributions_by_author/jones1998/) was reconstructed by *Jones et al.* [1998] using temperature sensitive paleoclimatic multi-proxy data from 10 sites worldwide that include tree rings, ice cores, corals, and historical documents. Only four of the proxy data series go back before 1400 AD and, therefore, data prior to about 600 years ago are more uncertain than temperature reconstructions after that time [see also *Jones et al.* 2001]. It is worth to note that a similar series extending through 1400-1998 A.D. was reconstructed by *Mann et al.* [1998]. An overview of important paleoclimate proxies and their uses is presented by *Stokstad* [2001].

The second example, shown in Figure 2 is one of the longest series of instrumental meteorological observations, the series of mean annual temperature at Paris/Le Bourget extending through 1764-1995. This is a typical example of a set of similar series of temperatures of European countries that go back to the 18th century (available from ftp.cru.uea.ac.uk).

The third example, shown in Figure 3, is the longest available record of river runoff in Greece, the Boeoticos Kefisos river runoff. The river is located to the north of Athens and the time series length is 91 years (hydrologic years 1907-08 to 1997-98).

A common characteristic in all three examples is that a local overyear average (plotted in the figures as 5- and 25-year average) is not stable but, rather, it exhibits significant variability. For example, in the first series (Figure 1) during the 16th century there is a falling trend of the local average, which is inverted during ca. 1650-1750, becomes again falling during ca. 1750-1850, and becomes rising thereafter. In the second series (Figure 2), there are falling trends during ca. 1765-1790, 1825-1875, and 1950-1980, and a rising trend during 1875-1950. (Note that the periods of falling and rising trends do not coincide with those of the first series; also, in similar series of different European towns the periods of falling and rising trends are different.) In the third series (Figure 3), there is an amplifying falling trend since 1920. (Notably, this trend agrees with a similar trend of the annual precipitation in the area; [*Nalbantis et al.*, 1993]).

A visual assessment of the magnitude of the overyear variability and trends can be done by comparing the actual time series with a series of white noise (independent, identically distributed variates). Such a synthetic series of white noise with length and marginal statistical characteristics equal to those of the first original series has been plotted in the middle panel of Figure 1. Clearly, the original series differs significantly, in a statistical sense, from white noise; the variability of the latter in the aggregation level of 25 years is much lower than that of the former.

The apparent falling and rising trends in all our example time series can be considered as climate changes or variations, but are they deterministic? This may be a difficult question as the definition of determinism belongs to the sphere of philosophy. Here we will attempt to give a practical answer based on *Borel's* [1920] view [see also *von Plato*, 1994, p. 41], who identifies determinism with predictability saying that there is indeterminism whenever Nature's actions are unpredictable. Clearly, then, these trends are only identified *a posteriori* (they cannot be predicted *a priori*); they are not regular, and it is difficult to attribute them to a cause-and-effect process. But, even if a deterministic model was available that could predict these trends accurately, this prediction would be done by operating the model in a fine timescale; the large-scale characteristics would be obtained by aggregating the results of the fine timescale. In this case, the complete time series would be deterministic and not merely the coarse timescale features, i.e. the overyear trends. Thus, the answer to the question set is negative.

Besides, there is not useful at all to consider the large-scale variability as a succession of deterministic trends. There would be, indeed, a reason if our purpose was to "subtract" these trends from the signal (a procedure known as detrending) so that the residual looks like the random signal of the middle panel of Figure 1. In fact, however, this would not help at all, because it is always our main objective to do projections for the future, and there is no reasonable way to do any projection of the "deterministic" part T_i. Again, the solution is to avoid decomposing of the original process X_i and treat it as a whole.

In conclusion, our interpretation is that the complete time series in all examples, including large-scale trends, must be regarded as stochastic signals. In this interpretation, the large-scale trends in the time series are closely related [e.g., *Evans*, 1996] to the well-known 'Hurst phenomenon', the special behavior of hydrologic and other geophysical time series discovered by *E. H. Hurst* [1951]. This behavior (also termed the 'Joseph effect' due to *Mandelbrot* [1977, p. 248]) has been characterized as the tendency of wet years to cluster into wet periods or of dry years to cluster into drought periods and, therefore, it has been also known as large-scale persistence. Equivalently, this behavior can be viewed as the result of falling and rising trends of the geophysical time series or, else, the large-scale variability of the time series. *Hurst* [1951] expressed mathematically his discovery in terms of a scaling exponent, H, which became known after him as the Hurst exponent (or coefficient). In a white noise series, such as the one plotted in the middle panel of Figure 1, H takes the value 0.5 whereas in all real-world time series H is greater than 0.5 and smaller than 1. The Hurst coefficient (the mathematics of which will be discussed later) provides a direct means to demonstrate the close relationship of large-scale variability with the Hurst phenomenon: if H is calculated using the original time series the result will be much higher than 0.5 (see section 4) but, if the same calculation is repeated for the detrended series, H will be around 0.5.

The above interpretation may seem similar (from a practical point of view) to that by *Klemes* [1974], who attributed the Hurst phenomenon to nonstationary means. However, there is a fundamental difference here. Nonstationarity of the mean would be the case if there existed a deterministic function expressing the mean as a function of time. On the contrary, we maintain that observed rising or falling trends are not deterministic components but rather large-scale stochastic fluctuations.

Consequently, a stochastic representation of hydrometeorological time series that is consistent with the varying climate hypothesis must be also consistent with the Hurst phenomenon. It is well-known that the most common stochastic models such as autoregressive (AR) models, moving average (MA) models, or combinations of the two (ARMA) are inappropriate to represent the Hurst phenomenon. However, several types of stochastic models able to reproduce the Hurst phenomenon have been proposed [Mandelbrot, 1965; *Mandelbrot and Wallis*, 1969a, b, c; *Mandelbrot*, 1971; Ditlevsen, 1971; Mejia et al., 1972; see also *Bras and Rodriguez-Iturbe*, 1985, pp. 210-280]. In addition, *Koutsoyiannis* [2000] has proposed a generalized framework for reproducing short- or large-scale persistence based on a generalized autocovariance function. Using the latter method, we have generated a second synthetic series, again with length and marginal statistical characteristics (now including the Hurst exponent), equal to those of the original series of temperature anomalies, which we have plotted in the lower panel of Figure 1. In this case, the large-scale variability agrees (in a statistical sense) with that of the original series (although periods of falling and rising trends are obviously different).

2.2 Basic assumptions and notation

Let X_i denote a hydrometeorological process with i = 1, 2, …, denoting discrete time. In the context of this paper we assume that X_i is not periodic, which means that our time scale is at least annual. Rather, we will assume that the process is stationary, a property that does not preclude it to exhibit large scale variability. Further, let us denote its mean $\mu := E[X_i]$, its autocovariance

$$\gamma_j := \mathrm{Cov}[X_i, X_{i+j}], \quad j = 0, \pm 01, \pm 2, \ldots \tag{2}$$

and its autocorrelation

$$\rho_j := \mathrm{Corr}[X_i, X_{i+j}] = \gamma_j / \gamma_0, \quad j = 0, \pm 1, \pm 2, \ldots \tag{3}$$

We also denote $\sigma := \sqrt{\gamma_0}$ the process standard deviation.

Let k be a positive integer that represents a timescale larger than the basic timescale of the process X_i. We denote $Z_i^{(k)}$ the aggregated stochastic process on that timescale, i.e.,

$$Z_i^{(k)} := \sum_{l=(i-1)k+1}^{ik} X_l \tag{4}$$

Obviously, for $k = 1$, $Z_i^{(1)} \equiv X_i$; for $k = 2$,

$$Z_1^{(2)} := X_1 + X_2, \quad Z_2^{(2)} := X_3 + X_4, \quad Z_3^{(2)} := X_5 + X_6, \ldots \tag{5}$$

for $k = 3$,

$$Z_1^{(3)} := X_1 + X_2 + X_3, \quad Z_2^{(3)} := X_4 + X_5 + X_6, \quad Z_3^{(3)} := X_7 + X_8 + X_9, \ldots \tag{6}$$

etc. The statistical characteristics of $Z_i^{(k)}$ for any timescale k can be derived from those of X_i. For example, the mean is

$$E[Z_i^{(k)}] = k\,\mu \tag{7}$$

whilst the variance and autocovariance (or autocorrelation) is more difficult to derive as it depends on the specific structure of γ_j (or ρ_j).

Thus, in the simplest case, where X_i is white noise (different X_i are independent identically distributed random variables), $\gamma_j = 0$ (and $\rho_j = 0$) for $j \neq 0$. Apparently then, the aggregated process will have variance

$$\gamma_0^{(k)} := \mathrm{Var}[Z_i^{(k)}] = k\,\gamma_0 \tag{8}$$

and autocovariance and autocorrelation

$$\gamma_j^{(k)} := \mathrm{Cov}[Z_i^{(k)}, Z_{i+j}^{(k)}] = 0, \quad \rho_j^{(k)} := \mathrm{Corr}[Z_i^{(k)}, Z_{i+j}^{(k)}] = 0, \quad \text{for } j \neq 0 \tag{9}$$

The hypothesis is set forward that hydrometeorological processes exhibit scale invariant properties at any scale longer than annual, i.e.,

$$(Z_i^{(k)} - k\,\mu) \stackrel{\mathrm{d}}{=} \left(\frac{k}{l}\right)^H (Z_j^{(l)} - l\,\mu) \tag{10}$$

where the symbol $\stackrel{d}{=}$ stands for equality in (finite dimensional joint) distribution and H is the Hurst exponent. Equation (10) is valid for any integer i and j (that is, the process is stationary) and any timescales k and l (≥ 1). If X_i is assumed Gaussian, this equation defines in discrete time the process known as fractional Gaussian noise (FGN), which was introduced by *Mandelbrot* [1965]. (In fact, the FGN process is typically defined in continuous time, e.g., *Saupe* [1988, p. 82]; *Abry et al.* [1995]; this however is not our scope here). It is easily shown [e.g., *Bras and Rodriguez-Iturbe*, 1985, p. 221] that the process defined by (10) reproduces the Hurst phenomenon. In our scope, we will avoid using the standard name 'fractional Gaussian noise' for the process defined by (10) for several reasons: The first term, fractional, is not easily understandable, unless combined with fractals, which is not necessary, when dealing with hydrologic statistics. The second term, Gaussian, may be not appropriate for several hydrologic processes that are asymmetric (non-Gaussian); there is no need to restrict our analysis to processes that are Gaussian. The third term, noise, usually describes a random and unstructured process, which is not the case in hydrologic processes that are structured. Instead we will use the term simple scaling stochastic process or simple scaling signal (SSS; can equivalently read as self similar signal).

As a consequence of (10), for $i = j = l = 1$ we get

$$\gamma_0^{(k)} := \mathrm{Var}[Z_i^{(k)}] = k^{2H} \gamma_0 \tag{11}$$

Thus, the standard deviation is a power law of the scale or level of aggregation k with exponent H. The extremely simple relation (11) can serve as the basis for estimating H [*Montanari et al.*, 1997; *Koutsoyiannis*, 2001b; see also section 3.3 below] thus avoiding the use of the original estimation technique by *Hurst* [1951] that is based on the so called rescaled range and also avoiding a number of problems relating with it [*Koutsoyiannis*, 2001b; see also section 3.3 below]. It is easy then to show [*Koutsoyiannis*, 2001b] that, for any aggregated timescale k, the autocorrelation function is independent of k, and given by

$$\rho_j^{(k)} = \rho_j = (1 / 2) [(j + 1)^{2H} + (j - 1)^{2H}] - j^{2H}, \quad j > 0 \tag{12}$$

We can characterize the SSS as a simplified model of reality, which uses one single parameter H to express the correlation structure of the process, noting that it is much more effective in representing hydrometeorological series than, for instance, the ARMA processes.

3. Statistical estimation and prediction under the SSS representation

In this section we will derive estimations for the most common statistics that are used in hydrologic estimation, prediction and testing under the hypothesis that the process of interest is SSS. We will assume that our sample is a time series of length n whose items correspond to consequent time instances, i.e., $X_1, \ldots, X_n$.

3.1 Estimation of mean

The simpler statistic estimated from a time series is the average $\overline{X}$ with estimator

$$\overline{X} := \frac{1}{n} \sum_{i=1}^{n} X_i \tag{13}$$

This is an unbiased estimator regardless of the type of the process X_i, i.e.,

$$E[\overline{X}] = \mu \tag{14}$$

In classic statistics, its variance is

$$\mathrm{Var}[\overline{X}] = \frac{\sigma^2}{n} \tag{15}$$

which, however, is not valid in SSS. Instead, observing that $\overline{X} = Z^{(n)}/n$ (where for simplicity we have omitted the subscript $i = 1$ of $Z_i^{(n)}$) and using (11) we obtain

$$\mathrm{Var}[\overline{X}] = \frac{\sigma^2}{n^{2-2H}} \tag{16}$$

We remind that the square root of Var[$\overline{X}$] is the standard error in estimating the true mean from the observed time series. For H = 1/2 both (15) and (16) result in the same standard error, which is inversely proportional to the square root of the length of the time series. For large H, however, the difference of (15) and (16) becomes very significant. To demonstrate this, we consider a time series of 100 years observations (n = 100) and assume H = 0.8. The classic statistics says that the estimation error is 1/10 of the process standard deviation. However, the correct standard error, as given by (16), is 1/2.5 of the process standard deviation, i.e., 4 times larger. Moreover, to have an estimation error equal to 1/10 of the process standard deviation, the required length of the time series would be 100 000 years! Obviously, this dramatic difference should induce substantial differences in other common statistics, as well, as we will see in the following sections.

3.2 Estimation of variance and standard deviation for known Hurst coefficient

The typical variance estimator

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \tag{17}$$

is no longer an unbiased estimator, as is in classic statistics. To show this, we rewrite (17) as

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} [(X_i - \mu) - (\overline{X} - \mu)]^2 \tag{18}$$

and, after algebraic manipulations, also considering (13), we get

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \mu)^2 - \frac{n}{n-1} (\overline{X} - \mu)^2 \tag{19}$$

Taking expected values in (19) and also considering (16) we obtain

$$E[S^2] = \frac{n - n^{2H-1}}{n-1} \sigma^2 \tag{20}$$

which proves that (17) is unbiased only when H = 0.5. Consequently, the SSS unbiased variance estimator for any known H is

$$\tilde{S}^2 := \frac{n-1}{n-n^{2H-1}} S^2 = \frac{1}{n\,(1-n^{2H-2})} \sum_{i=1}^{n} (X_i - \bar{X})^2 \tag{21}$$

This expression has some similarity with earlier expressions that corrected the variance estimator in terms of the lag 1 autocorrelation [*Matalas*, 1967; *O'Connell*, 1974; *Salas*, 1993, p. 19.11]. We can then consider $\tilde{S}$ (the square root of $\tilde{S}^2$) as an approximately unbiased estimator of the standard deviation σ. The variance of the estimator of S for a normal distribution of X in the classic statistics is [e.g., *Yevjevich*, 1972, p. 193]

$$\mathrm{Var}[S] \approx \frac{\sigma^2}{2(n-1)} \tag{22}$$

In the SSS case it is very difficult to derive the variance of estimator S or $\tilde{S}$ in an analytical manner. Instead we performed a systematic Monte Carlo study (again for a normal distribution of X), from which we concluded that

$$\mathrm{Var}[\tilde{S}] \approx \frac{\sigma^2}{\kappa(H)\, n^{\lambda(H)}\, (1-n^{2H-2})} \tag{23}$$

where

$$\kappa(H) := \begin{cases} [0.5\,(1-H)]^{-0.5} & H \le 0.6 \\ [1.2\,(1-H)]^{-1.1} & H > 0.6 \end{cases} \tag{24}$$

and

$$\lambda(H) := \begin{cases} 1 & H \le 0.6 \\ [2.5\,(1-H)]^{0.45} & H > 0.6 \end{cases} \tag{25}$$

It is easily verified that, when H = 0.5, (23) shifts to (22).

To demonstrate the consequences of using the inappropriate classic estimators of variance and standard deviation we have performed a Monte Carlo experiment. We generated a long series of SSS with $H = 0.8$, $\mu = 2$ and $\sigma = 0.5$. From this series we constructed an ensemble of 100 samples each with length $n = 100$ or 50 and we estimated the sample standard deviations using both estimators. We did the same for aggregation levels $k = 1$ to 10 when $n = 100$ and $k = 1$ to 5 when $n = 50$ (so that in any aggregation level the number of items n / k is at least 10). The results are shown graphically in Figure 4 in a logarithmic plot of standard deviation versus scale. The true standard deviation for each aggregation level k is obtained from (11). Its empirical values are obtained as the averages of the 100 samples. We observe that, at the basic scale ($k = 1$), the classic estimators underestimate the true standard deviation by about 6% and 9% for $n = 100$ and 50, respectively. The percentage of underestimation increases to about 20% at the largest scale used, i.e., $k = n / 10$. However, the SSS estimates agree perfectly with the theoretical curve (the two curves are practically indistinguishable).

This increasingly underestimated standard deviation with the increase of the scale k, when the classic estimator is used, has another important consequence: it obscures the presence of the Hurst phenomenon for small samples. Specifically, in the logarithmic plot of Figure 4 we observe that the slope of the curve of classic estimate of standard deviation versus scale is not constant, as implied by the scaling law (11), but decreases with the increase of scale k. Not only does it result in underestimation of the Hurst coefficient itself, but it may also lead us to consider that this slope tends to 0.5 for large scales, in which case we will reject the presence of the Hurst phenomenon.

In addition, the standard deviation, over the 100 samples, of the sample standard deviation at the basic scale are 0.067 is 0.043 for $n = 50$ and 100, respectively; the classic estimator (22) predicts 0.051 and 0.036, respectively, whereas estimator (23) predicts the values 0.061 and 0.046, respectively, which are much better than the classic ones.

3.3 Simultaneous estimation of variance and Hurst coefficient

When the Hurst exponent is unknown, which is the most usual case when dealing with an observed time series, (21) cannot be applied as it contains the unknown *H*. Traditionally, the estimation of *H* has been based on the original Hurst's algorithm which is based on the concept of the so called rescaled range, a statistic with many difficulties and inaccuracy [*Koutsoyiannis*, 2001b; see also discussion below]. Other algorithms have been studied by *Montanari et al.* [1997]. Here we propose a new algorithm, which is consistent with the SSS statistics. This algorithm is based on standard deviations s_k for timescales k ranging from 1 to a maximum value $k' := n / 10$. This maximum value was chosen so that s_k can be estimated from at least 10 data values.

Combining (20) and (11) we get

$$E[S_k] \approx c_k(H)\, k^H \sigma \tag{26}$$

where

$$c_k(H) := \sqrt{\frac{n/k - (n/k)^{2H-1}}{n/k - 1}} \tag{27}$$

Based on (26) we can estimate simultaneously H and σ in terms of minimising a fitting error. Among several expressions that were tried for this error, denoted e, the following has been proved (using Monte Carlo experiments) to be the most efficient (i.e., to have the narrowest confidence intervals)

$$e^2(\sigma, H) := \sum_{k=1}^{k'} \frac{\{\ln E[S_k] - \ln s_k\}^2}{k} = \sum_{k=1}^{k'} \frac{[\ln \sigma + H \ln k + \ln c_k(H) - \ln s_k]^2}{k} \tag{28}$$

Taking the derivatives of e^2 with respect to $\ln \sigma$ and H end equating to zero we get

$$\frac{1}{2}\frac{\partial e^2(\sigma, H)}{\partial \ln \sigma} = \ln \sigma \sum_{k=1}^{k'} \frac{1}{k} + H \sum_{k=1}^{k'} \frac{\ln k}{k} + \sum_{k=1}^{k'} \frac{\ln c_k(H)}{k} - \sum_{k=1}^{k'} \frac{\ln s_k}{k} = 0 \tag{29}$$

$$\frac{1}{2}\frac{\partial e^2(\sigma, H)}{\partial H} = \ln \sigma \sum_{k=1}^{k'} \frac{d_k(H)}{k} + H \sum_{k=1}^{k'} \frac{d_k(H) \ln k}{k} + \sum_{k=1}^{k'} \frac{d_k(H) \ln c_k(H)}{k} - \sum_{k=1}^{k'} \frac{d_k(H) \ln s_k}{k} = 0 \quad (30)$$

where

$$d_k(H) := \ln k + \frac{\partial \ln c_k(H)}{\partial H} = \ln k - \frac{(n/k)^{2H-1} \ln(n/k)}{(n/k - 1)\, c_k(H)^2} \quad (31)$$

Eliminating $\ln \sigma$ we get

$$H = \left\{ \sum_{k=1}^{k'} \frac{d_k(H) \ln s_k}{k} \sum_{k=1}^{k'} \frac{1}{k} - \sum_{k=1}^{k'} \frac{d_k(H) \ln c_k(H)}{k} \sum_{k=1}^{k'} \frac{1}{k} - \sum_{k=1}^{k'} \frac{\ln s_k}{k} \sum_{k=1}^{k'} \frac{d_k(H)}{k} \right.$$
$$\left. + \sum_{k=1}^{k'} \frac{\ln c_k(H)}{k} \sum_{k=1}^{k'} \frac{d_k(H)}{k} \right\} \Big/ \left\{ \sum_{k=1}^{k'} \frac{d_k(H) \ln k}{k} \sum_{k=1}^{k'} \frac{1}{k} - \sum_{k=1}^{k'} \frac{\ln k}{k} \sum_{k=1}^{k'} \frac{d_k(H)}{k} \right\} \quad (32)$$

In this equation H appears in both sides. However, it can be easily solved in an iterative manner. Assuming an initial value H = 0.5 and substituting it in the right-hand side we directly calculate (in the left-hand side) an improved estimate and we continue this way. The method converges quickly. Having found H, σ is obtained directly from (29). Alternatively, it can be estimated from (21) using the standard deviation of the finest time scale only.

To demonstrate the performance of the algorithm we have done a Monte Carlo experiment using the ensemble of 100 samples already discussed in section 3.2 with n = 50 and 100, generated for H = 0.80 and σ = 0.50. For comparison we also used similar series of white noise (H = 0.50). The resulting values of H and σ using the above algorithm are shown in Figure 5 by means of box plots. The figure shows a good performance of the algorithm, especially for n = 100, whereas for n = 50 and H = 0.8 a slight underestimation of both H and σ appears. For comparison we have also applied a much simpler algorithm, which is based on the classic estimator of σ (equation (17); see *Montanari et al.*, [1997]; *Koutsoyiannis* [2001b]), which, as shown in Figure 5, apparently results in underestimation of both H and σ even for n = 100. Finally, another comparison is performed using the traditional Hurst's algorithm based on the rescaled range for the ensemble with n = 100 and H = 0.8. Clearly, Figure 5 shows that this algorithm is inappropriate. Its exhibits a negative bias, which is rather slight, but, more importantly, the dispersion of the results is more than double that of the proposed algorithm. Notably, in a non-ignorable percentage of samples, this algorithm resulted in H greater than 1, which is mathematically inconsistent, as well in H smaller than 0.5, which is physically inconsistent (although mathematically acceptable). On the contrary, the proposed algorithm never resulted in H greater than 1; also, it never resulted in H smaller than 0.5 when the true H is 0.8. Some additional information on the behavior of the proposed algorithm is provided in Figure 6, where the estimated standard deviation is plotted against the estimated Hurst exponent. The figure indicates that for H < 0.8 the two statistics are practically uncorrelated but for higher H they become positively correlated.

3.4 Estimation of distribution quantiles

The problem of estimation of quantiles of hydrologic variables, for certain values of probability of exceedance or non-exceedance, is central to hydrologic statistics. Clearly, the above analyses lead to dramatic differences of these quantiles in comparison with their classic estimations.

To demonstrate this we consider the case where the variable X is normally distributed. In this case the classic estimator of the u-quantile (the value of the variable for probability of non-exceedance u) is

$$\hat{X}_u = \overline{X} + z_u\, S \tag{33}$$

where z_u is the u-quantile of the standard normal distribution. Its classic confidence limits and for confidence coefficient γ are given by [e.g., *Stedinger et al.*, 1993, p. 18.30]

$$\hat{x}_{u_{1,2}} = \hat{x}_u \pm z_{(1\,+\,\gamma\,/2)}\, \varepsilon_u \tag{34}$$

where

$$\varepsilon_u = \frac{s}{\sqrt{n}} \sqrt{1 + \frac{z_u^2}{2}} \tag{35}$$

For the SSS case, assuming a known H, the estimator of the u-quantile for any scale k becomes

$$\hat{Z}_u^{(k)} = k\,\bar{X} + z_u\,k^H\,\tilde{S} \tag{36}$$

The confidence limits can be estimated using (16) and (23) and also assuming that, when X is normally distributed, $\bar{X}$ and $\tilde{S}$ are stochastically independent (as in the classic case), a fact verified by Monte Carlo simulations. After algebraic manipulations, these confidence limits become

$$\hat{z}_{u_{1,2}}^{(k)} = \hat{z}_u^{(k)} \pm z_{(1+\gamma/2)}\,\varepsilon_u \tag{37}$$

with

$$\varepsilon_u = k\,\frac{s}{n^{1-H}} \sqrt{1 + \frac{z_u^2}{\kappa(H)\,n^{\lambda(H)}\,(n/k)^{2H-2}\,(1 - n^{2H-2})}} \tag{38}$$

To demonstrate the difference of the classic and SSS estimators we have plotted in Figure 7 the 95% confidence limits of both cases using the first synthetic sample of the ensemble already described in section 3.2 for n = 100. We remind that the parameters used to generate this synthetic sample are H = 0.8, μ = 2 and σ = 0.5. This figure is for the basic timescale (k = 1). We observe in Figure 7 that the point estimations of quantiles for u ranging from 0.01 to 0.99 differ only slightly due to the small departure of the SSS standard deviation from the classic one. However, the confidence interval determined by the SSS estimator is 3 to 4 times wider than in that of the classic estimator. Interestingly, the true (theoretical) quantiles, which are also shown in Figure 7, lie outside of the 95% classic confidence limits for the entire probability domain shown in the graph. On the contrary, it lies within the 95% SSS confidence limits, again for the entire probability domain. This demonstrates the inappropriateness of classic estimators and the appropriateness of the SSS ones.

When both the standard deviation and the Hurst exponent are unknown, apparently the confidence intervals will be even wider. Their theoretical determination, however is very difficult and therefore Monte Carlo methods [e.g. *Ripley*, 1987] must be the appropriate choice.

3.5 Estimation of cross-covariances and cross-correlations

Assuming that two processes X_i and Y_i are both SSS with common H and mutually correlated, the typical covariance estimator

$$S_{XY} := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) \tag{39}$$

is again a biased estimator. Reasoning as above, and also assuming, by analogy to (11), that the aggregated covariance is

$$\mathrm{Cov}[X_1 + \ldots + X_n, Y_1 + \ldots + Y_n] = n^{2H}\, \mathrm{Cov}[X_1, Y_1] \tag{40}$$

we conclude that the unbiased estimator for any known H is

$$\widetilde{S}_{XY} := \frac{n-1}{n - n^{2H-1}} S_{XY} = \frac{1}{n - n^{2H-1}} \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) \tag{41}$$

Here we observe that the correlation coefficient is

$$R_{XY} := \frac{S_{XY}}{S_X S_Y} = \frac{\widetilde{S}_{XY}}{\widetilde{S}_X \widetilde{S}_Y} \tag{42}$$

Thus, the classic estimator of the (cross-) correlation coefficient remains valid also for SSS.

This is demonstrated in Figure 8. As in section 3.2, we performed here another Monte Carlo experiment by generating bivariate synthetic samples with common $H = 0.8$ and length

n = 100 or 50. The other characteristic parameters were μ = 2 and 3, and σ = 0.5 and 1.2 for the first and second variable, respectively. The theoretical cross-correlation coefficient was 0.85. To generate the bivariate synthetic samples we followed the multivariate method by *Koutsoyiannis* [2000]. From the ensembles of 100 series we estimated the empirical cross-correlations for all timescales, which, as shown in Figure 8, agree well with the theoretical expectation.

3.6 Estimation of auto-covariances and auto-correlations

In the case of autocorrelation coefficients the situation is different as it has been shown that for series with nonzero autocorrelation, the typical estimator of autocovariance is biased downward [*Wallis and O'Connell*, 1972; *Salas*, 1993, p. 19.10]. The typical estimator of the lag l autocovariance is [e.g., *Salas*, 1993, p. 19.10]

$$G_l := \frac{1}{n} \sum_{i=1}^{n-l} (X_i - \bar{X})(X_{i+l} - \bar{X}) \tag{43}$$

This can be written as

$$G_l = \frac{1}{n} \sum_{i=1}^{n-l} [(X_i - \mu) - (\bar{X} - \mu)]\ [(X_{i+l} - \mu) - (\bar{X} - \mu)] \tag{44}$$

After algebraic manipulations, also considering (13), we get

$$G_l = \frac{1}{n} \sum_{i=1}^{n-l} (X_i - \mu)\ (X_{i+l} - \mu) - \frac{1}{n} (\bar{X} - \mu) \sum_{i=1}^{n-l} [(X_i - \mu) + (X_{i+l} - \mu)] + (\bar{X} - \mu)^2 \tag{45}$$

Assuming that l is small in comparison with n so that we can interchange $n - l$ and n, and also extend the second sum of (45) over all i, we obtain

$$G_l \approx \frac{1}{n} \sum_{i=1}^{n-l} (X_i - \mu)\ (X_{i+l} - \mu) - (\bar{X} - \mu)^2 \tag{46}$$

Taking expected values (and again ignoring the difference of $n - l$ and n), we find that

$$E[G_l] \approx \gamma_l - \frac{\sigma^2}{n^{2-2H}} \tag{47}$$

This means that an approximately unbiased estimator of γ_l will be

$$\tilde{G}_l := G_l + \frac{1}{n^{2-2H}} \tilde{S}^2 = G_k + \frac{n-1}{n^{3-2H} - n} S^2 \tag{48}$$

Consequently, an approximately unbiased estimator of the autocorrelation coefficient ρ_k will be

$$\tilde{R}_l := \frac{\tilde{G}_l}{\tilde{S}^2} = R_l\left(1 - \frac{1}{n^{2-2H}}\right) + \frac{1}{n^{2-2H}} \tag{49}$$

where R_l is the classic estimator of the autocorrelation coefficient, i.e.,

$$R_l := \frac{n}{n-1} \frac{G_k}{S^2} \tag{50}$$

Clearly, the classic estimator of autocorrelation is biased and the bias becomes very high when H is high. This is demonstrated in Figure 9 by means of the Monte Carlo experiment already discussed in section 3.2. Not only is the classic estimation of autocorrelation coefficient significantly lower than the theoretical value but it also vanishes off (becomes practically zero) for lags 5-10 thus obscuring the long-term persistence of the process. This may have dramatic consequences as the process may be taken as a short-memory one and be modeled using typical ARMA models, which of course are inappropriate. On the contrary, the SSS estimator developed above captures very well the long-term persistence of the process and agrees perfectly with the theoretical autocorrelation function.

4. Case studies

In this section we discuss further the three real-world time series that were introduced in section 2.1. In Figure 10 we have plotted the standard deviation of the Jones's proxy time series of temperature anomalies versus timescale. Clearly, the standard deviation of the aggregated process is a power function of timescale and this is apparent even using the classic estimator of standard deviation. The exclusion of the data of the last century, which can be suspect for anthropogenic influence, does not change the shape of the curve of standard deviation versus slope, as also shown in Figure 10. The slope of this curve in the log-log plot is high and differs significantly from 0.5, the value that characterizes white noise. Using the method described in section 3.3 we estimated the Hurst coefficient, which is the slope of this curve, at 0.92. Had the classic estimation of standard deviation been used, H would be 0.86, a value that corresponds to the mean slope of the classic standard deviation versus scale.

In the upper panel of Figure 11 we have plotted the lag-1 and lag-2 autocorrelation function versus scale. The SSS model implies that these autocorrelations are independent of scale (equation (12)). The classic empirical estimations of the autocorrelation do not verify this theoretical expectation as for large scales the autocorrelation decreases. However, the SSS estimations of autocorrelation agree well with the model, as they are almost constant even for scales as large as 50 years. In the lower panel of Figure 11 we have plotted the autocorrelation versus lag for the basic scale ($k = 1$). Due to the large length of the time series, the long-term persistence of the time series is obvious here even when the classic autocorrelation estimators are used. However, in this case the empirical autocorrelation depart from theoretical ones for lags > 40, even when the value $H = 0.86$ is used. When the SSS estimation of autocorrelation is used along with the value $H = 0.92$, the empirical autocorrelation function agrees perfectly with the model.

In Figure 12 we have plotted the point estimates and the 99% confidence limits of the quantiles of the temperature anomalies for probability of non-exceedance u ranging from 1% to 99%. This is done for two timescales, the basic time scale ($k = 1$) that represents the annual variation of temperature anomaly, and the 30-year time scale, which typically is assumed as sufficient to smooth out the annual variations and be a representative value of the climate. (For the latter we have used the averaged rather than aggregated time series, i.e., $z^{(30)}/30$). We observe in Figure 12 that the variation of the 30-year average, is only slightly lower than that of the annual values. For example, the 99%-quantile of annual temperature anomaly slightly exceeds 0.6^{o}C above average, whereas the 99%-quantile of the thirty-year average is almost 0.5^{o}C above average. These values refer to the point estimations of quantiles. If we consider the upper 99% confidence limits of quantiles, these values become about 1.0^{o}C and 0.9^{o}C above average, respectively. These figures indicate that an increase of the thirty-year average temperature by 0.5^{o}C, does not provide strong statistical evidence of an unusual change of climate. (We note that the observed temperature increase since 1850 is around 0.5^{o}C).

The results related to this example, however, contain a high degree of uncertainty due to the proxy character of the time series. This problem however, does not emerge in our next example, the Paris temperature time series. In Figure 13 we have plotted the standard deviation of Paris temperature versus timescale. Again, the standard deviation of the aggregated process is a power function of timescale and this is apparent even using the classic estimator of standard deviation. The slope of this curve in the log-log plot, which represents the Hurst coefficient, estimated using the method described in section 3.3, is 0.81 (it would be 0.79 if the classic estimation of standard deviation was used). The autocorrelation coefficients shown in Figure 14 also verify the presence of long-term persistence, the appropriateness of the proposed SSS estimator of autocorrelation, and the large departure of the classic estimations from SSS estimators. Otherwise, as shown in Figure 2, there is no remarkable pattern in this time series that would require further statistical analysis.

An interesting pattern exists in our third example time series, the Boeoticos Kifisos runoff, shown in Figure 3, which as already discussed in section 2.1, exhibits a falling trend since 1920, lasting 78 out of a total of 91 years. The characteristic plot of standard deviation versus time scale, shown in Figure 15, again verifies the presence of long-term persistence. As in the previous examples, the relation of standard deviation versus scale is a power law; the exponent is 0.78.

Typically, in hydrologic statistics a trend is detected using the Kendall's τ statistic [e.g. *Kottegoda*, 1980, p. 32] defined as

$$\tau := \frac{4\,p}{n\,(n-1)} - 1 \tag{51}$$

where p is the number pairs of observations (x_j, x_i; $j > i$) in which $x_j < x_i$. In a random series τ has mean 0, variance $2(2n + 5)/9n(n - 1)$, and distribution converging rapidly to normal. In our example, the application of the test results in $\tau = 0.40$ for $n = 78$. The standard deviation of τ is 0.077, and eventually Kendall's test results in rejection of the null hypothesis that a trend does not exist for an attained significance level as low as 8.8×10^{-8}; this is typically considered as sufficient statistical evidence that a trend really exits. However, this is not correct because the time series is not random but an SSS series. To find the value of the standard deviation of τ we can now use stochastic simulation. Thus, we generated an ensemble of 100 time series with $n = 78$ and $H = 0.78$, from which we found that the standard deviation of τ is 0.173 and the attained significance level of the Kendall's test becomes now 0.01. Still, this is not absolutely correct because it does not correspond exactly to the formulation of the null and alternative hypotheses. We recall that, to formulate the hypotheses to be tested, we did an exploratory data analysis of the complete 91-year series and we located this trend to 78 out of the 91 years. On the other hand, it is known that the validity of the confirmatory tests is based on the assumption that the investigator developed the hypothesis prior to examining the data [*Hirsch et al.*, 1993, p. 17.5]. To re-establish the validity of the test, we performed a different stochastic simulation that is consistent with the procedure of formulating the hypotheses. Specifically, we generated an ensemble of 100 time series, each with $n = 91$ and $H = 0.78$, and in each of these series we located that 78-year period which gave the maximum value of τ (in absolute value). Now the standard deviation of τ over the 100 series is 0.252 and the attained significance level of the Kendall's test is 0.055. This means that the trend is not statistically significant at the 1% or even the 5% significance level.

The falling trend could be alternatively viewed as a downward jump. This jump, as shown in Figure 16, can be located between 1971 and 1972. The 65-year period before the jump has an average of 439.3 mm and the 26-year period after the jump has an average of 276.6 mm, whereas the entire 91-year average is 392.8 mm. In classic statistics, the difference 493.3 – 276.6 = 162.7 mm would be tested for being statistically significant using the typical statistical test for equality of means. Indeed, this test results in rejection of the hypothesis of equality of the means at an attained significance level as low as 8.2×10^{-6}. This typical test, however, is incorrect because it was based on the classic statistics. To perform a more accurate test we performed stochastic simulation. In each of the 100 generated 91-year long synthetic series already discussed in the previous paragraph we located the 26-year period with the minimum average and we took the difference from the average of the entire 91-year series. Then we determined the probability that this difference exceeds 392.8 – 276.6 = 116.2 mm, which is as high as 22%. This probability is the attained significance level of the test and this means that the hypothesis of equality of means is not rejected at the usual significant levels.

In conclusion no statistically significant trend or jump is detected in the Boeoticos Kifisos runoff time series. The above statistical analysis defeats earlier analyses of the same time series [e.g. *Nalbantis et al.*, 1993], which detected statistically significant trends or jumps using classic statistics.

5. Summary and conclusions

Despite of the intensive research of the recent years on climate change, according to the experts of climate modeling, “the current state of affairs is not satisfactory” [*Barnett et al.*, 1999] in terms of prediction capabilities of the climate evolution and quantification of the related uncertainty. Moreover, the unpredictability of future climate in deterministic terms may by a structural characteristic of the climate system (rather than a matter of current weaknesses of models) since, according to *von Storch et al.* [2001], "climate must be considered as a stochastic system, and our climate simulation models as random number generators".

Therefore, probability-based, statistic or stochastic methods may be good alternatives to quantify uncertainty, even under a varying climate. However, hydrologic statistics, the branch of hydrology that deals with uncertainty, has been based on the implicit assumption of a stable climate. This disagrees with the fact that climate has ever, through the planet history, changed irregularly on all time scales, a fact that becomes obvious from long hydroclimatic time series. Observed shifts in such time series were often regarded as deterministic components (trends or jumps) and removed from the time series so that the residual can be processed using classic statistics. This, however, is not the correct way as the shifts are in fact stochastic rather than deterministic. A stochastic basis for dealing with these shifts is offered by the Hurst phenomenon, which, in fact, is no more than the simple scaling behavior of the variation of hydroclimatic quantities with scale. The Hurst coefficient is the exponent of the power-law relationship between the aggregated standard deviation at any timescale and the scale length.

When we employed the simple scaling hypothesis and revisited the typical statistical descriptors used in hydrologic statistics under this hypothesis we discovered that

1. the classic sample average remains an unbiased estimator of the true mean, but its variance, which expresses the uncertainty of the its estimation, is dramatically higher than the value given by the classic statistics;
2. the classic estimator of variance is no more unbiased and the variance thereof is different from that of the classic statistics;
3. the quantiles of a given distribution function differ from those of the classic statistics and their confidence intervals are radically wider than those implied by the classic statistics;

4. the classic estimators of cross-correlations between two variables remain almost unbiased, but those of autocorrelations are highly biased.

For all above statistical descriptors, new generalized estimators are determined which are unbiased or almost unbiased under the simple scaling hypothesis. These estimators depend on the Hurst exponent H, in addition to the other dependencies used in classic statistics. The classic estimators are derived as special cases of the generalized estimators when $H = 0.5$. The Hurst exponent itself has been considered as another statistic and an algorithm has been developed to estimate it, avoiding the use of the concept of rescaled range.

The application of the developed statistical framework to three hydrometeorological time series with lengths ranging from 91 to 992 years showed that all three series agree perfectly with the scaling hypothesis (Hurst phenomenon). In addition, it is shown that several patterns within these times series would be regarded as evident trends or shifts if classic statistical tests were used, but using modified tests based on the scaling hypothesis it turns out that they are nothing more than regular behavior of the time series.

In conclusion, the analyses of this paper cast a warning that the classic hydrologic statistics describes only a portion of the natural uncertainty of hydroclimatic processes, because it is based on the implicit assumption of a stable climate and, in addition, its use may characterize a regular behavior of hydroclimatic processes as unusual phenomenon. Furthermore, the analyses show that it is feasible to adapt the classic hydrologic statistics so as to quantify the total uncertainty under a varying climate. Obviously, further and more detailed analyses of several related issues of hydrologic statistics, and investigations of a large number of data sets, are needed before a concrete base of methodologies, appropriate for different types of water resources problems, can be established.

Acknowledgments. The research leading to this paper was performed within the framework of the project *Modernization of the supervision and management of the water resource system of Athens*, funded by the Water Supply and Sewage Corporation of Athens. The author wishes to thank the directors of EYDAP and the members of the project committee for the support of the research. The discussions with Andreas Efstratiadis that stimulated many of the analyses of this paper are gratefully acknowledged.

References

Abry, P., P. Gonçalvés and P. Flandrin, Wavelets, spectrum analysis and 1/*f* processes, in *Wavelets and Statistics*, edited by A. Antoniadis and G. Oppenheim, Springer-Verlag, New York, 1995.

Barnett, T.P., K. Hasselmann, M. Chelliah, T. Delworth, G. Hegerl, P. Jones, E. Rasmusson, E. Roeckner, C. Ropelewski, B. Santer, and S. Tett, Detection and attribution of recent climate change: A status report, *Bulletin of the American Meteorological Society*, 80, 2631-2659, 1999.

Borel, E., Radioactivité, probabilité, déterminisme, Oeuvres de Emile Borel, 4, 2189-2196, 1920.

Bras, R.L. and I. Rodriguez-Iturbe, *Random Functions in Hydrology*, Addison-Wesley, 1985.

Chow, V. T., Statistical and probabilistic analysis of hydrologic data, in *Handbook of Hydrology*, edited by V. T. Chow, McGraw-Hill, New York, 1964.

Ditlevsen, O. D., Extremes and first passage times, Doctoral dissertation presented to the Technical University of Denmark, Lyngby, Denmark, 1971.

Evans, T. E., The effects of changes in the world hydrological cycle on availability of water resources, Chapter 2 in *Global Climate Change and Agricultural Production: Direct and Indirect Effects of Changing Hydrological, Pedological and Plant Physiological Processes*, edited by F. Bazzaz and W. Sombroek, Food and Agriculture Organization of the United Nations and John Wiley, Chichester, 1996.

Gardner, W. A., *Introduction to Random Processes with Applications to Signals and Systems*, McGraw-Hill, New York, 1989.

Haan, C. T., *Statistical Methods in Hydrology*, The Iowa State University Press, USA, 1977.

Hirsch, R. M., D. R. Helsel, T. A. Cohn, and E. J. Gilroy, Statistical analysis of hydrologic data, in *Handbook of Hydrology*, edited by D. R. Maidment, McGraw-Hill, 1993.

Hurst, H. E., Long term storage capacities of reservoirs, *Trans. ASCE*, 116, 776-808, 1951.

Jones, P. D., K. R. Briffa, T. P. Barnett, and S. F. B. Tett, High-resolution paleoclimatic records for the last millennium: Interpretation, integration and comparison with General Circulation Model control-run temperatures, *Holocene, 8 (4)*, 455-471, 1998.

Jones, P. D., T. J. Osborn and K. R. Briffa, The evolution of climate over the last millennium, *Science*, 292(5517), 662-667, 27 Apr 2001.

Kite, G. W., *Frequency and Risk Analyses in Hydrology*, Water Resources Publications, Littleton, Colorado, 1988.

Klemes, V., The Hurst phenomenon: A puzzle?, *Water Resour. Res.*, 10(4) 675-688, 1974.

Kottegoda, N. T., *Stochastic Water Resources Technology*, Macmillan Press, London, 1980.

Koutsoyiannis, D., A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research*, 36(6), 1519-1534, 2000.

Koutsoyiannis, D., Coupling stochastic models of different time scales, *Water Resources Research*, 37(2), 379-392, 2001a.

Koutsoyiannis, D., Hurst phenomenon and fractional Gaussian noise made easy, *J. of Hydrol.*, submitted, 2001b.

Ledley, T. S., E. T. Sundquist, S. E. Schwartz, D. K. Hall, J. D. Fellows, and T. L. Killeen, Climate change and greenhouse gases, *EOS*, 80(39), 453, 1999.

Mandelbrot, B. B., Une classe de processus stochastiques homothetiques a soi: Application a la loi climatologique de H. E. Hurst, *Compte Rendus Academie Science*, 260, 3284-3277, 1965.

Mandelbrot, B. B., A fast fractional Gaussian noise generator, *Water Resour. Res.*, 7(3), 543-553, 1971.

Mandelbrot, B. B., *The Fractal Geometry of Nature*, Freeman, New York, 1977.

Mandelbrot, B. B., and J. R. Wallis, Computer experiments with fractional Gaussian noises, Part 1, Averages and variances, *Water Resour. Res.*, 5(1), 228-241, 1969a.

Mandelbrot, B. B., and J. R. Wallis, Computer experiments with fractional Gaussian noises, Part 2, Rescaled ranges and spectra, *Water Resour. Res.*, 5(1), 242-259, 1969b.

Mandelbrot, B. B., and J. R. Wallis,. Computer experiments with fractional Gaussian noises, Part 3, Mathematical appendix, *Water Resour. Res.*, 5(1), 260-267, 1969c.

Mann, M. E. R. S. Bradley, and M. K. Hughes, Global-scale temperature patterns and climate forcing over the past six centuries, *Nature* 392, 779-787, 1998.

Matalas, N. C., Mathematical assessment of synthetic hydrology, *Water Resour. Res.*, 3(4), 937-945, 1967.

Mejia, J. M., I. Rodriguez-Iturbe and D. R. Dawdy, Streamflow simulation, 2, The broken line process as a potential model for hydrologic simulation, *Water Resour. Res.*, 8(4), 931-941, 1972.

Montanari, A., R. Rosso snd M. S. Taqqu, Fractionally differenced ARIMA models applied to hydrologic time series, *Water Resour. Res.*, 33(5), 1035-1044, 1997.

Nalbantis, I., N. Mamassis, and D. Koutsoyiannis, Le phénomène récent de sécheresse persistante et l' alimentation en eau de la cite d' Athènes, *Publications de l'Association Internationale de Climatologie, 6eme Colloque International de Climatologie,* éditeur P. Maheras, Thessaloniki, Septembre 1993, 6, 123-132, Association Internationale de Climatologie, Aix-en-Provence Cedex, France, 1993.

National Research Council, (1991). Committee on Opportunities in the Hydrologic Sciences, *Opportunities in the Hydrologic Sciences*, National Academy Press, Washington, DC.

O'Connell, P. E., Stochastic modelling of long-term persistence in streamflow sequences, PhD dissertation, Imperial College of Science and Technology, London, 1974.

Przybylak, R., Temporal and spatial variation of surface air temperature over the period of instrumental observations in the Arctic, *International Journal of Climatology*, 20, 587-614, 2000.

Ripley, B. D., *Stochastic Simulation*, Wiley, New York, 1987.

Salas, J. D., Analysis and modeling of hydrologic time series, *Handbook of hydrology*, edited by D. Maidment, Chapter 19, pp. 19.1-19.72, McGraw-Hill, New York, 1993.

Saupe, D., Algorithms for random fractals, Chapter 2 in *The Science of Fractal Images*, edited by H.-O. Peitgen and D. Saupe, Springer-Verlag, 1988.

Shaw, E. M., *Hydrology in Practice*, 3rd edition, Chapman and Hall, London, 1994.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou, Frequency analysis of extreme events, Chapter 18 in *Handbook of Hydrology*, edited by D. R. Maidment, McGraw-Hill, 1993.

Stokstad, E. Myriad ways to reconstruct past climate, *Science*, 292(5517), 658-659, 27 Apr 2001.

Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, J. F. B. Mitchell, and G. J. Jenkins, External control of 20th century temperature by natural and anthropogenic forcings, *Science*, 290, 2133-2137, 2000.

von Plato, J., *Creating Modern Probability*, Cambridge University Press, Cambridge, 1994.

von Storch, H., J-S. von Storch and P. Müller, Noise in the climate system – ubiquitous, constitutive and concealing, in *Mathematics Unlimited – 2001and Beyond*, edited by B. Engquist and W. Schmid, Springer, Berlin, 2001.

Wallis, J. R., and P. E. O'Connell, Small sample estimation of ρ_1, *Water Resour. Res.*, 8(3), 707-712, 1972.

Yevjevich, V., *Probability and Statistics in Hydrology*, Water Resources Publications, Fort Collins, Colorado, 1972.

List of Figures

Figure 7 Point estimates of quantiles at the basic timescale (k = 1) and 95% confidence limits thereof estimated from a synthetic time series with length n = 100 generated with theoretical parameters $H = 0.8$, $\mu = 2$ and $\sigma = 0.5$.

Figure 8 Comparison of theoretical and empirical cross-correlation coefficients of a bivariate aggregated process $\mathbf{Z}_i^{(k)}$ versus timescale k for a Monte Carlo experiment with theoretical Hurst coefficient 0.8 for both variates, standard deviations 0.5 and 1.2 for the first and second variate, respectively, and theoretical cross-correlation coefficient 0.85.

Figure 9 Comparison of theoretical and empirical autocorrelation functions at the basic scale ($k = 1$) for a Monte Carlo experiment with theoretical $H = 0.8$ and $\sigma = 0.5$.

Figure 10 Standard deviation of the aggregated processes versus timescale (logarithmic plot) for the Jones's time series of the North Hemisphere temperature anomalies. For comparison we have also plotted the theoretical curve of the white noise model.

Figure 11 Autocorrelation coefficients of the Jones's time series of the North Hemisphere temperature anomalies: (up) lag 1 and lag 2 autocorrelations of the aggregated process versus timescale, k; (down) autocorrelation versus lag for the basic timescale, $k = 1$.

Figure 12 Point estimates of quantiles at the basic timescale (annual values, $k = 1$) and the 30-year timescale (30-year averages, $k = 30$), and 99% confidence limits thereof for the Jones's time series of the North Hemisphere temperature anomalies.

Figure 13 Standard deviation of the aggregated processes versus timescale (logarithmic plot) for the time series of mean annual temperature at Paris/Le Bourget. For comparison we have also plotted the theoretical curve of the white noise model.

Figure 14 Autocorrelation coefficients of the time series of mean annual temperature at Paris/Le Bourget: (up) lag 1 and lag 2 autocorrelations of the aggregated process versus timescale, k; (down) autocorrelation versus lag for the basic timescale, $k = 1$.

Figure 15 Standard deviation of the aggregated processes versus timescale (logarithmic plot) for the runoff time series of the Boeotikos Kifisos river basin.

Figure 16 Auxiliary sketch for testing the hypothesis of a jump at the runoff time series of the Boeotikos Kifisos river basin: plots of the original time series and the averages before and after the jump.

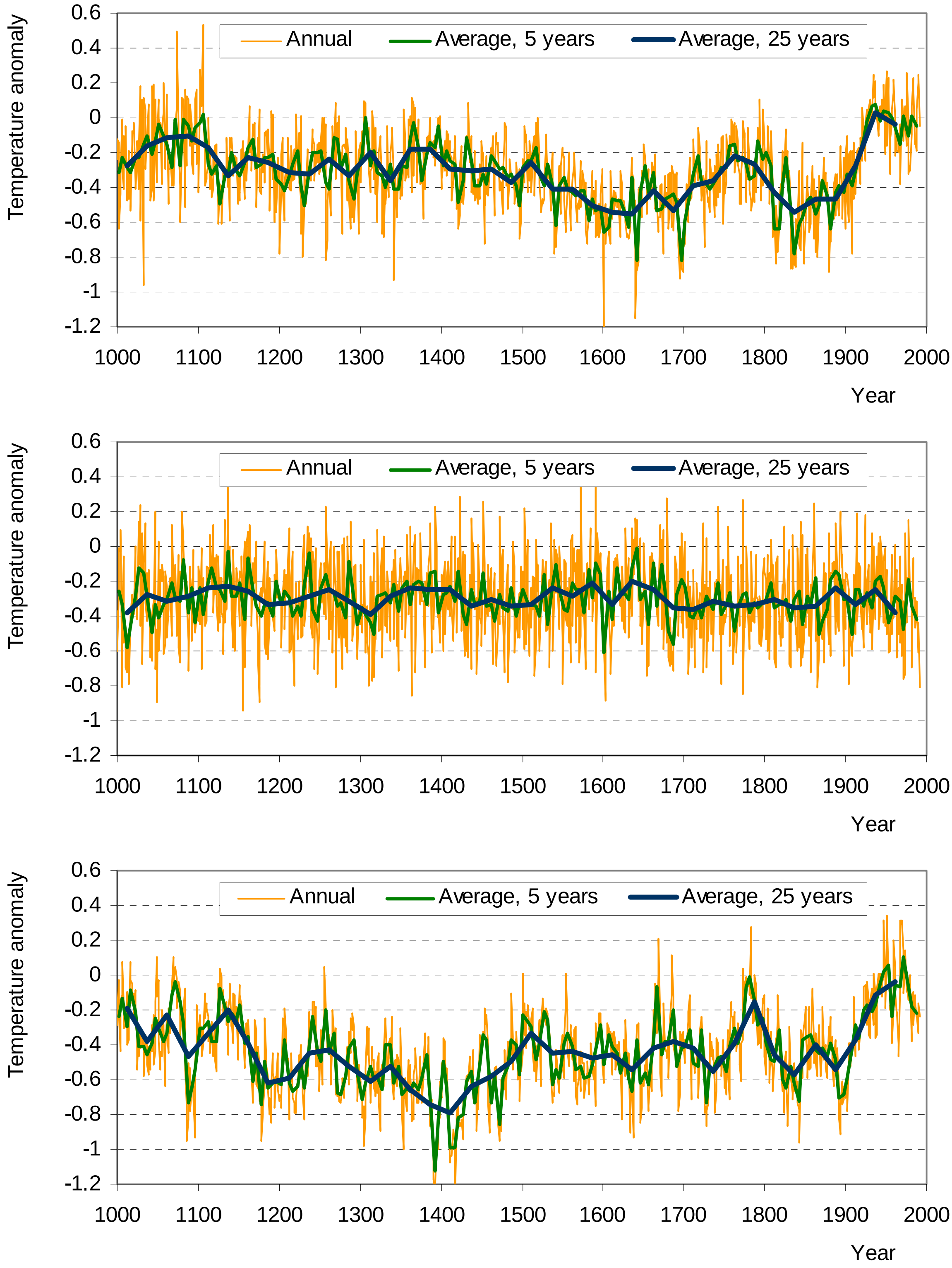


Figure 1 Plot of the North Hemisphere temperature anomalies, reconstructed by *Jones et al.* [1998] using proxy data (up). For comparison we have also plotted (middle) a series of white noise with statistics equal to those of the original series and (down) a synthetic simple scaling series with statistics equal to those of the original series.

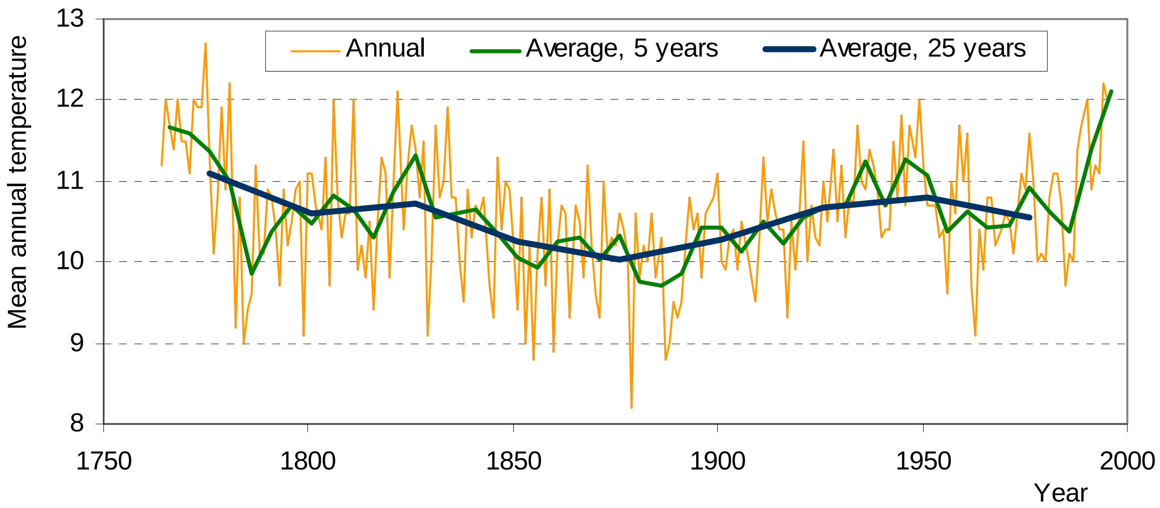


Figure 2 Plot of the time series of mean annual temperature at Paris/Le Bourget.

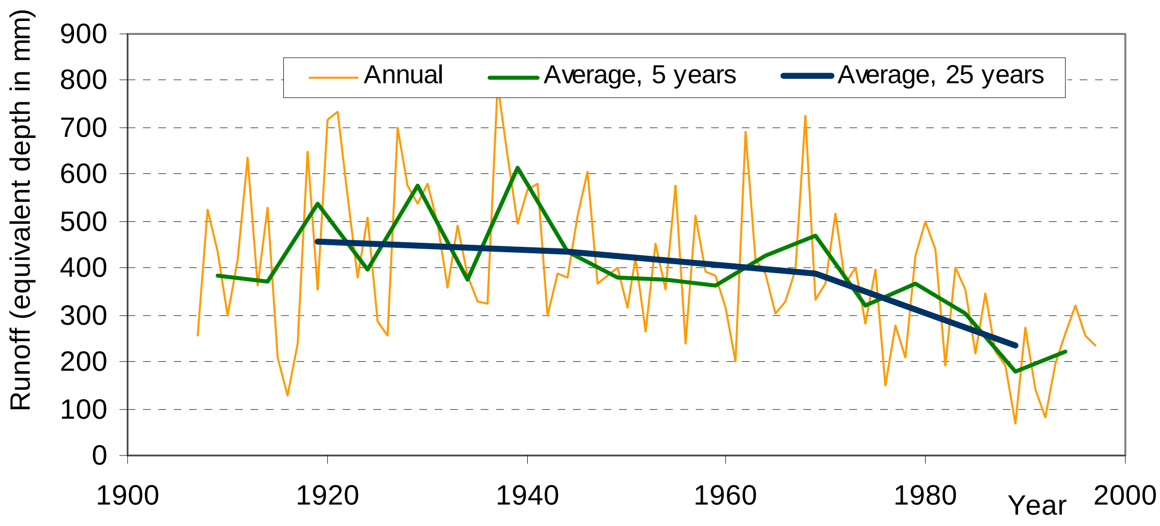


Figure 3 Plot of the time series of the equivalent runoff depth of the Boeotikos Kifisos river basin, Greece.

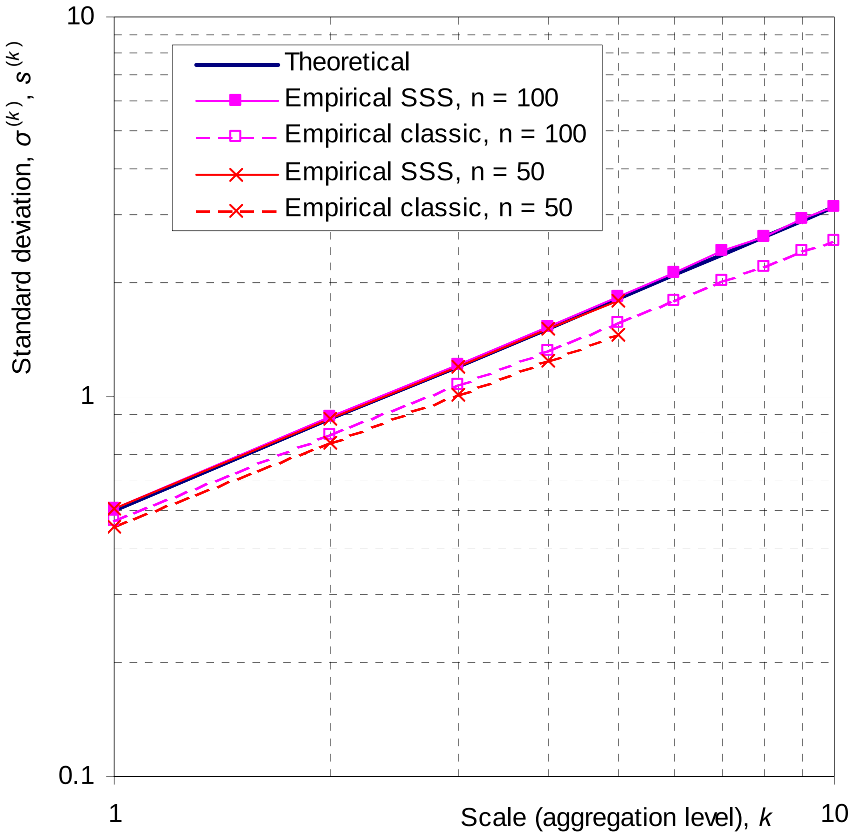
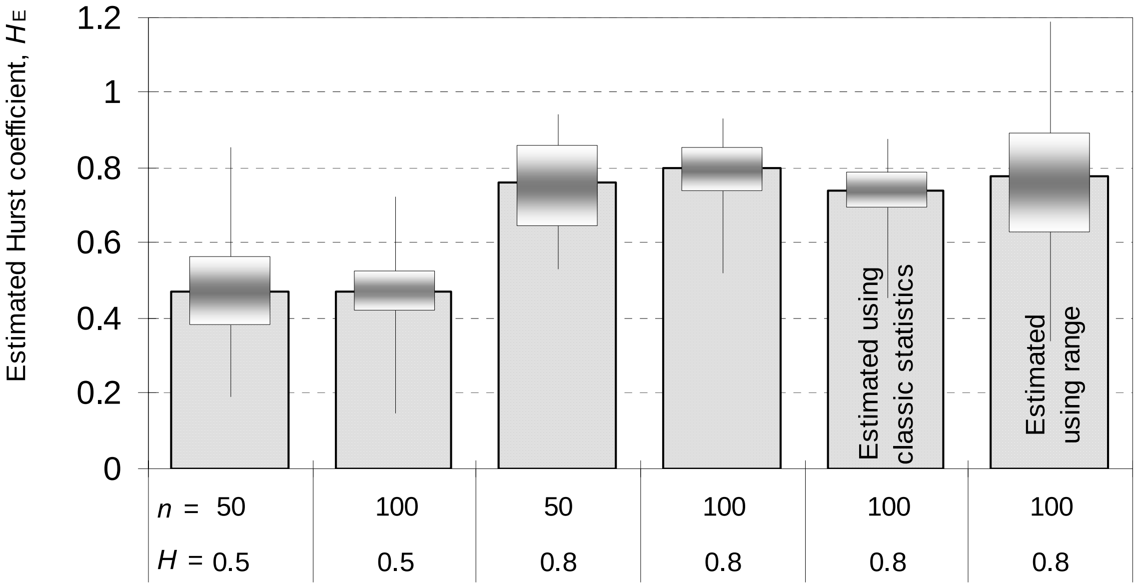


Figure 4 Comparison of theoretical and empirical standard deviation of the aggregated processes $Z_i^{(k)}$ versus timescale k (logarithmic plots) for a Monte Carlo experiment with theoretical $H = 0.8$ and $\sigma = 0.5$.



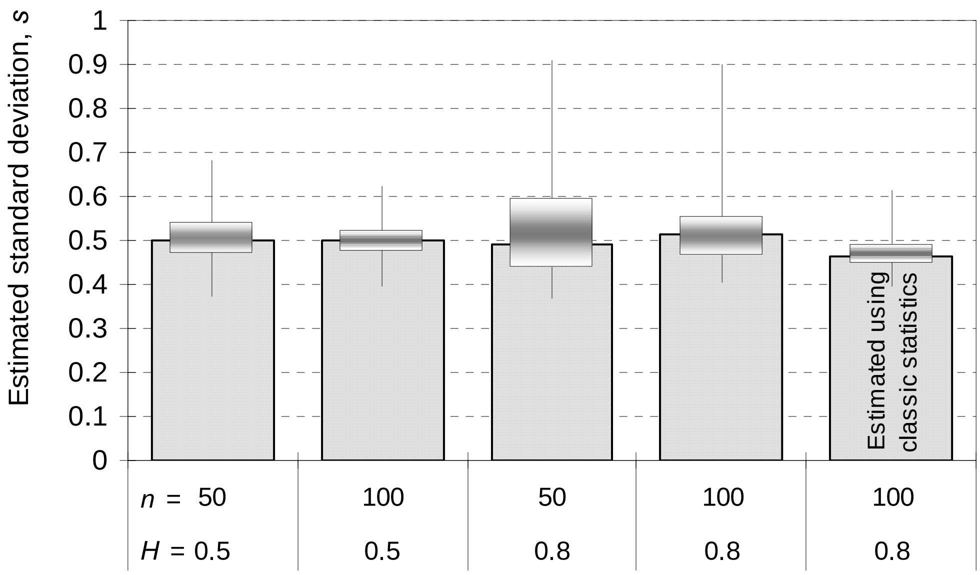


Figure 5 Box plots of the estimated Hurst coefficients (up) and standard deviations (down) from ensembles of synthetic series with theoretical H = 0.5 or 0.8 and σ = 0.5. Bars correspond to the median of 100 estimations (from 100 synthetic series), upper and lower edges of boxes correspond to the 75%- and 25%-quantiles, respectively, and whiskers correspond to the maximum and minimum estimated values. The first four boxes correspond to the proposed method (section 3.3), the fifth boxes correspond to estimates using classic statistics, and the sixth box in the upper panel corresponds to the estimation using the original Hurst's algorithm based on the rescaled range.

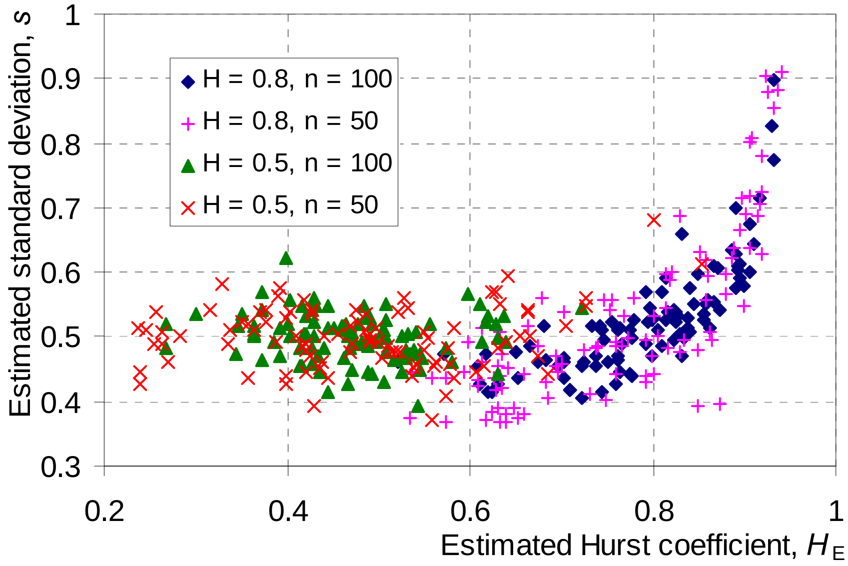


Figure 6 Estimated Hurst coefficients versus estimated standard deviations from the ensembles of synthetic series of the Monte Carlo experiment of Figure 5 and for the proposed estimation method (section 3.3).

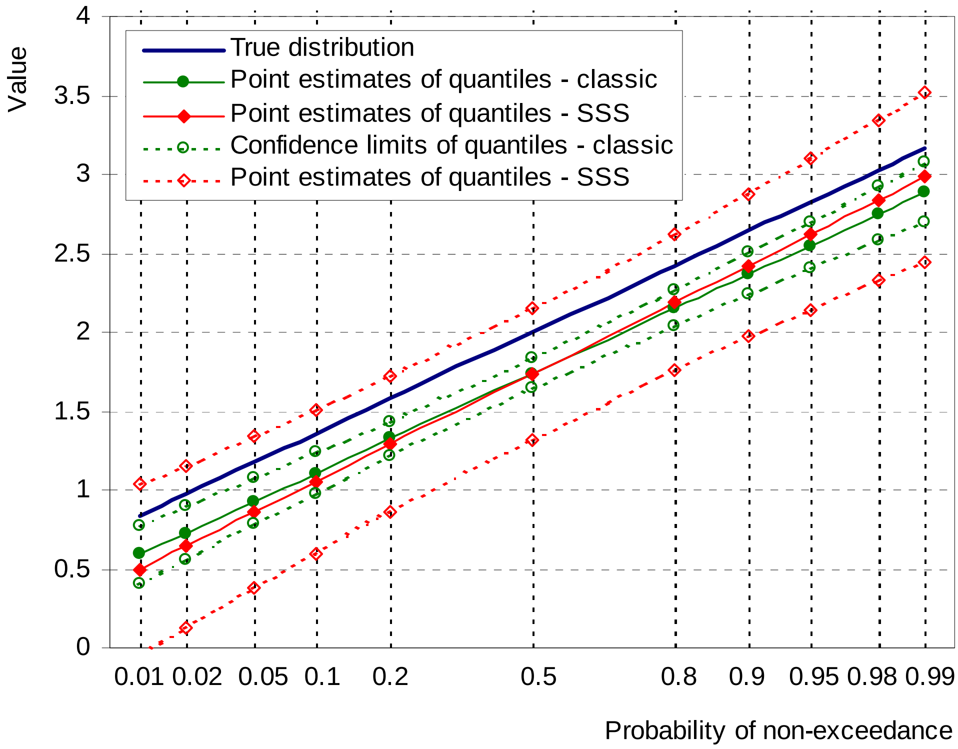


Figure 7 Point estimates of quantiles at the basic timescale (k = 1) and 95% confidence limits thereof estimated from a synthetic time series with length n = 100 generated with theoretical parameters $H = 0.8$, $\mu = 2$ and $\sigma = 0.5$.

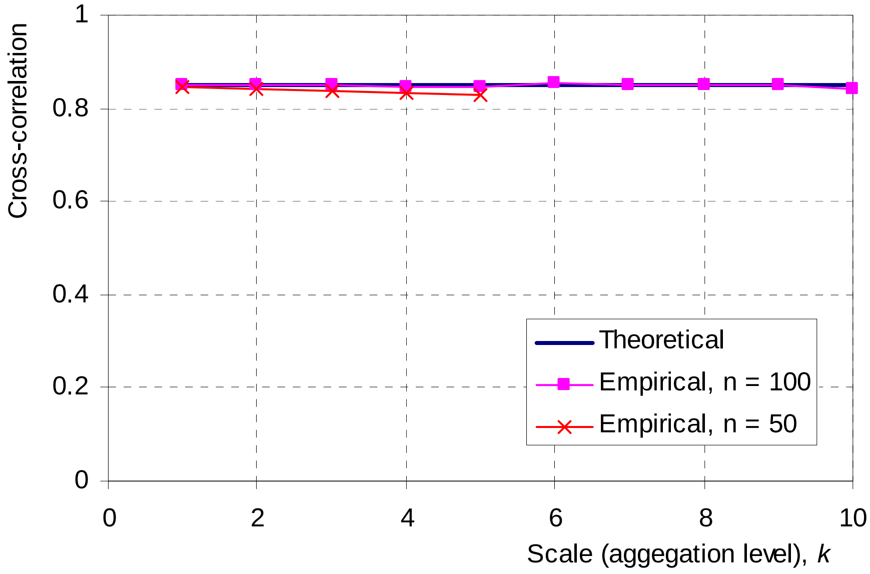


Figure 8 Comparison of theoretical and empirical cross-correlation coefficients of a bivariate aggregated process $\mathbf{Z}_i^{(k)}$ versus timescale k for a Monte Carlo experiment with theoretical Hurst coefficient 0.8 for both variates, standard deviations 0.5 and 1.2 for the first and second variate, respectively, and theoretical cross-correlation coefficient 0.85.

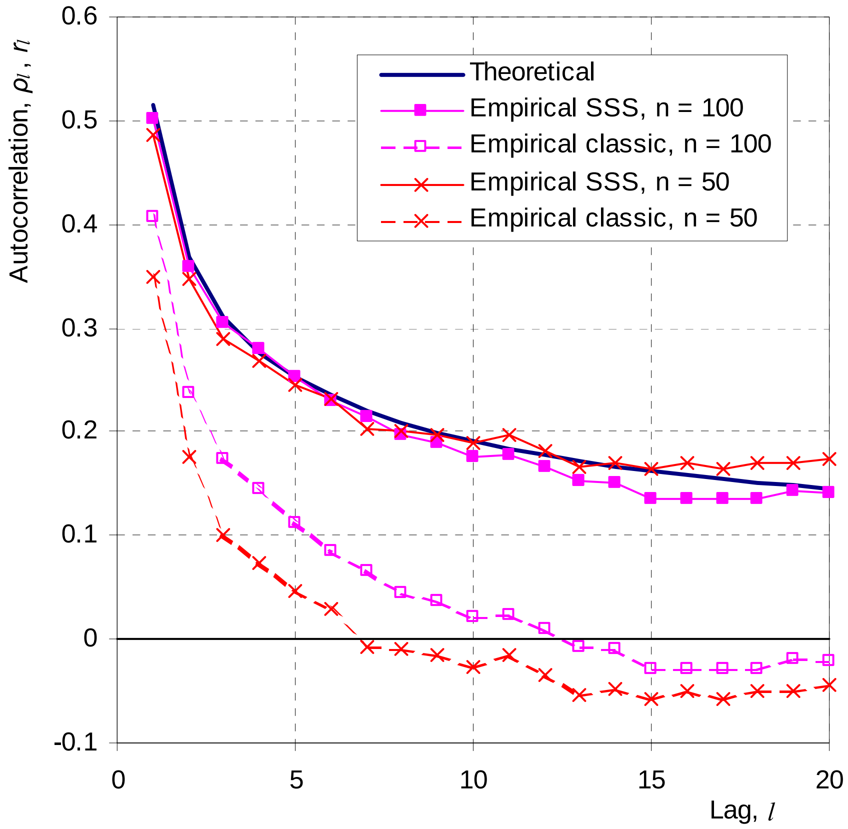


Figure 9 Comparison of theoretical and empirical autocorrelation functions at the basic scale ($k = 1$) for a Monte Carlo experiment with theoretical $H = 0.8$ and $\sigma = 0.5$.

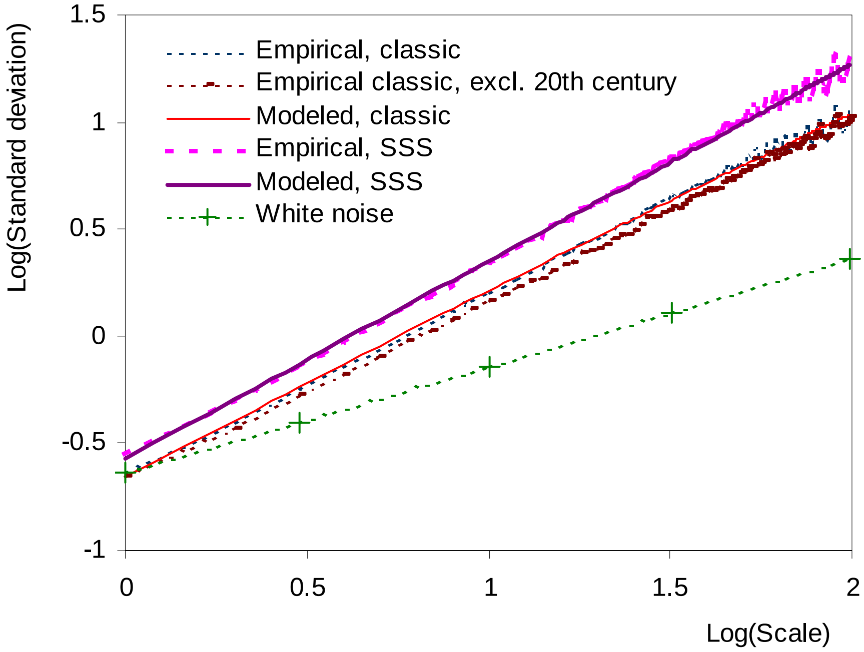
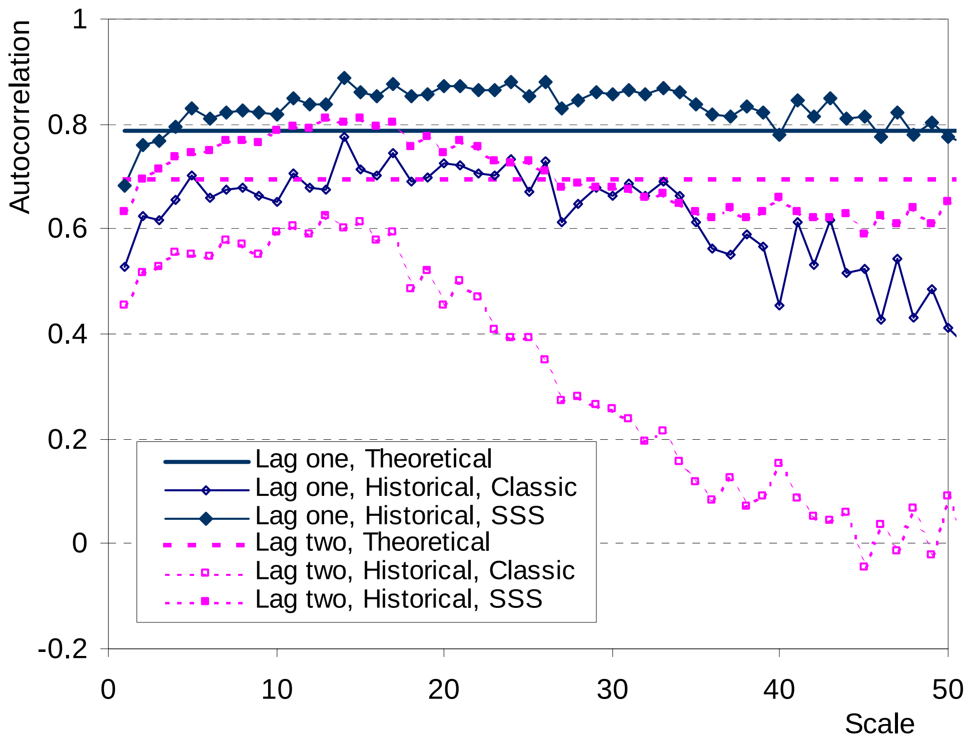


Figure 10 Standard deviation of the aggregated processes versus timescale (logarithmic plot) for the Jones's time series of the North Hemisphere temperature anomalies. For comparison we have also plotted the theoretical curve of the white noise model.



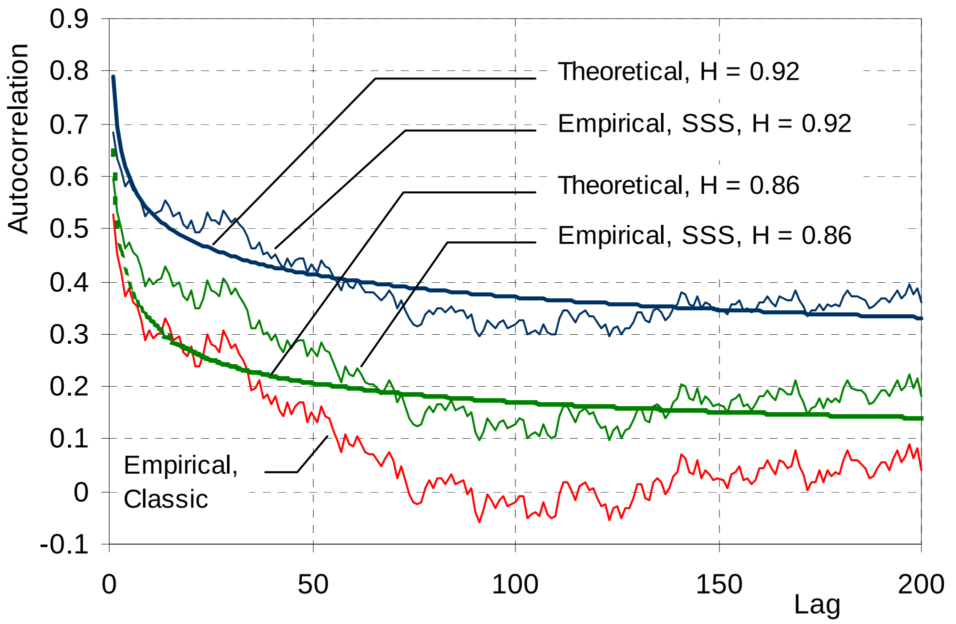


Figure 11 Autocorrelation coefficients of the Jones's time series of the North Hemisphere temperature anomalies: (up) lag 1 and lag 2 autocorrelations of the aggregated process versus timescale, k; (down) autocorrelation versus lag for the basic timescale, $k = 1$.

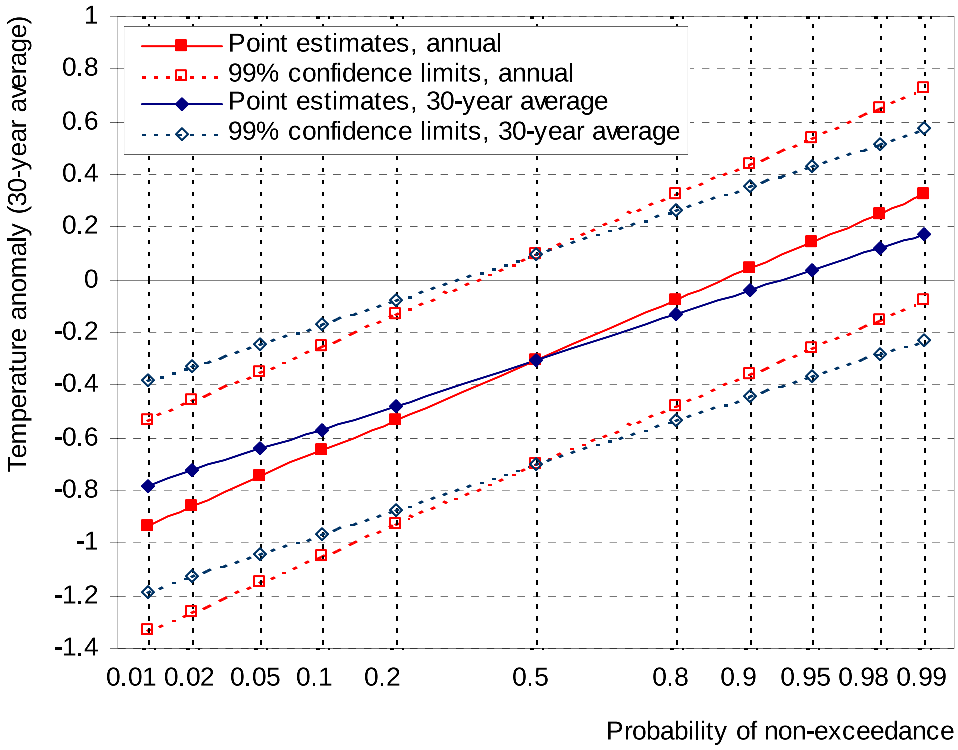


Figure 12 Point estimates of quantiles at the basic timescale (annual values, $k = 1$) and the 30-year timescale (30-year averages, $k = 30$), and 99% confidence limits thereof for the Jones's time series of the North Hemisphere temperature anomalies.

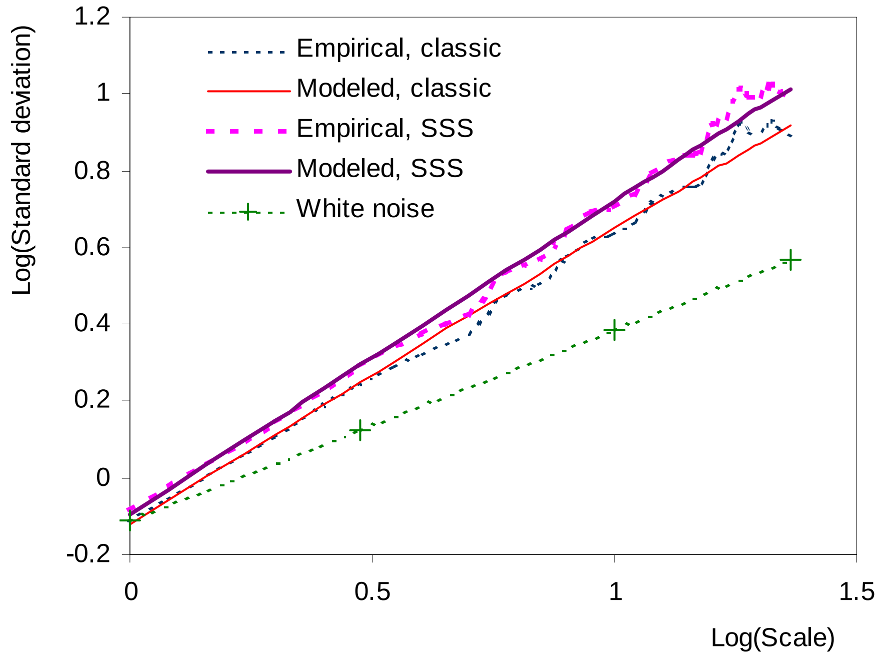


Figure 13 Standard deviation of the aggregated processes versus timescale (logarithmic plot) for the time series of mean annual temperature at Paris/Le Bourget. For comparison we have also plotted the theoretical curve of the white noise model.

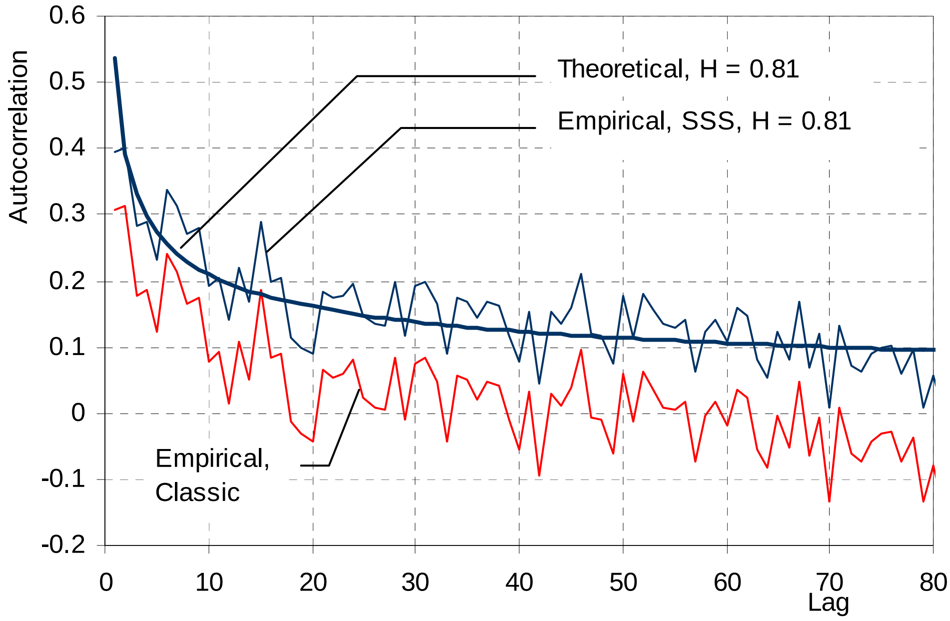


Figure 14 Autocorrelation coefficients of the time series of mean annual temperature at Paris/Le Bourget: (up) lag 1 and lag 2 autocorrelations of the aggregated process versus timescale, k; (down) autocorrelation versus lag for the basic timescale, $k = 1$.

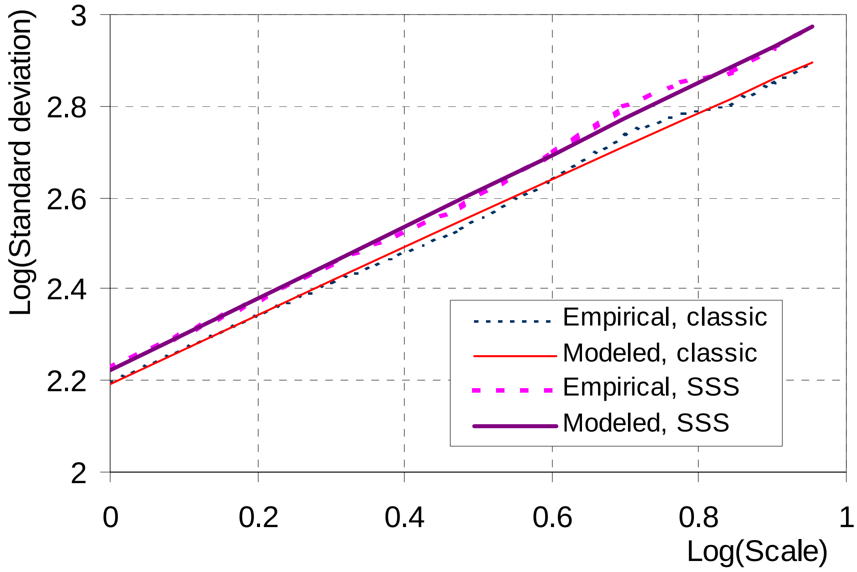


Figure 15 Standard deviation of the aggregated processes versus timescale (logarithmic plot) for the runoff time series of the Boeotikos Kifisos river basin.

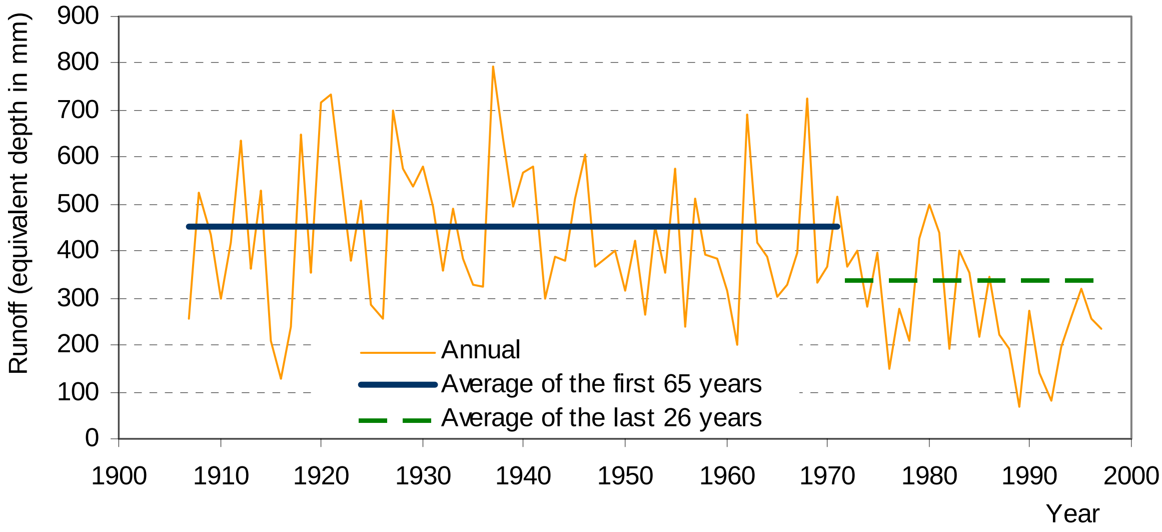


Figure 16 Auxiliary sketch for testing the hypothesis of a jump at the runoff time series of the Boeotikos Kifisos river basin: plots of the original time series and the averages before and after the jump.