Climate change, the Hurst phenomenon, and hydrological statistics

DEMETRIS KOUTSOYIANNIS

Department of Water Resources, School of Civil Engineering, National Technical University, Athens Heroon Polytechneiou 5, GR-157 80 Zographou, Greece

dk@itia.ntua.gr

Abstract The intensive research of recent years on climate change has led to the strong conclusion that climate has always, throughout the Earth's history, changed irregularly on all time scales. Climate changes are closely related to the Hurst phenomenon, which has been detected in many long hydroclimatic time series and is stochastically equivalent to a simple scaling behaviour of climate variability over time scale. The climate variability, anthropogenic or natural, increases the uncertainty of the hydrological processes. It is shown that hydrological statistics, the branch of hydrology that deals with uncertainty, in its current state is not consistent with the varying character of climate. 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 of a varying climate following a simple scaling law, and new estimators are studied which, in many cases, differ dramatically from the classical ones. 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 classical statistics.

Key words climate change; Hurst phenomenon; hydrological persistence; hydrological statistics; hydrological estimation; hydrological prediction; statistical testing; uncertainty

Changement climatique, phénomène de Hurst et statistiques hydrologiques

Résumé La recherche intensive des années récentes sur le changement climatique a conduit à la conclusion sûre que le climat a toujours changé dans l'histoire de la planète, et ceci de manière irrégulière à toutes les échelles de temps. Les changements climatiques sont étroitement liés au phénomène de Hurst, qui a été détecté dans de nombreuses séries temporelles longues d'hydroclimatologie et qui est stochastiquement équivalent à un comportement d'échelle simple de la variabilité climatique sur l'échelle de temps. La variabilité climatique, qu'elle soit d'origine anthropique ou naturelle, augmente l'incertitude liée aux processus hydrologiques. Il est démontré que l'hydrologie statistique, la branche de l'hydrologie qui s'occupe de l'incertitude, n'est pas, dans son état actuel, consistante avec le caractère variable du climat. Quelques caractéristiques statistiques typiquement utilisées en hydrologie comme les moyennes, les variances, les auto-corrélations et corrélations croisées, et le coefficient de Hurst, ainsi que leur variabilité, sont ré-examinées sous l'hypothèse d'un climat variable suivant une loi d'échelle simple. De plus de nouveaux estimateurs, pour la plupart très différents des estimateurs classiques, sont étudiés. Le nouveau cadre statistique est appliqué à des exemples réels, pour des travaux typiques comme l'estimation et le test d'hypothèses, où, à nouveau, les résultats diffèrent significativement de ceux des statistiques classiques.

Mots clefs changement climatique; phénomène de Hurst; persistance hydrologique; statistiques hydrologiques; estimation hydrologique; prédiction hydrologique; tests statistiques; incertitude

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, and on modelling the climate at the global scale using the so-called

climatic, or general circulation models. Climatic models describe some of the mechanisms of climate variability that are well understood, such as ice–albedo feedback, CO_2 cycles and greenhouse effects, ocean deep-water circulation, ocean–atmosphere interactions, land–atmosphere interactions, etc. They are capable of reproducing the large-scale distributions of pressure, temperature, precipitation and ocean-surface heat flux, and resemble sea-surface temperature anomalies related to the El Niño Southern Oscillation phenomena (e.g. Ledley *et al.*, 1999). Another important field of recent research is the detection and attribution of changes in the past climate. This has also been the subject of scientific debate on whether existing climatic records indicate a significant change of the present climate *vs* the past, and on whether detected changes are attributed to natural or anthropogenic forcings.

Thus, there is a number of studies detecting global warming in the past two decades and attributing them 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 (the latter especially for the last 35 years).

On the other hand, to invoke another recent study, Przybylak (2000) studies mean monthly temperatures of Arctic and sub-Arctic areas. The analyses show that in the Arctic, since the mid-1970s, the annual temperature shows no clear trend and the level of temperature in Greenland in the last 10–20 years is similar to that observed in the 19th century. This does not agree with predictions produced by some numerical climate models.

Barnett *et al.* (1999) state that, at present, it is not possible to distinguish the relative contributions of specific natural and anthropogenic forcings to observed climate change. 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" and conclude that "the current state of affairs is not satisfactory". If this is the case when dealing with the past and present climate, how can one trust 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 future the situation is different. In another comprehensive study by 36 climate modellers (Davey *et al.*, 2002), who tested the ability of 23 dynamical ocean–atmosphere models to correctly simulate fields of tropical sea-surface temperature, surface wind stress, and upper ocean vertically-averaged temperature, several model deficiencies are found, for example the facts that "interannual variability is commonly too weak in the models" and "no single model is consistent with observed behaviour in the tropical ocean regions".

Such deficiencies are related to the fact that climatic models are necessarily simplified representations of the complex climate system as they do not describe completely the dynamics of all involved processes. 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). Besides, 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, volcanic activity, etc.). Some of these factors are extremely difficult or even impossible to incorporate in models and to predict. Overall, as von Storch *et al.* (2001) put it, "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, it is well known that climate "changes irregularly, for unknown reasons, on all timescales" (National Research Council, 1991, p. 21).

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

In the last two decades, hydrologists have developed different strategies to deal with climate change (Lettenmaier *et al.*, 1996). One of these consists of the so-called downscaling of climatic model results into the area and time scale of interest. Another option is the adoption of prescribed climate change scenarios with plausible shifts in the average hydrological regime of the area of interest (e.g. changes 0 to +4°C in temperature and 0 to 25% in precipitation; Lettenmaier *et al.*, 1996). The first method suffers from accumulation of uncertainties and errors in the chain of models used, which potentially lead to poor representation of reality. For example, in a recent study by Carpenter & Georgakakos (2001) the large-scale climatic model used explains less than 20% of the observed precipitation variance and, even worse, results in significant scale bias (model precipitation even 5–25 times smaller than the actual one depending on the choice of the neighbouring model grid node, as displayed in their Fig. 6). The second method lacks quantification of uncertainty.

Traditionally, however, hydrological 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 hydrological statistics (e.g. Chow, 1964; Yevjevich, 1972; Kottegoda, 1980; Hirsch *et al.*, 1993; Stedinger *et al.*, 1993) has been well developed and applied in hydrological engineering studies. Hydrological statistics, though, has been based on the implicit assumption of a stable climate. The questions arise then: (a) Is hydrological statistics, in its present state, consistent with the assumption of a varying climate? (b) If not, what adaptations are needed to achieve this consistency? and (c) Can hydrological statistics be used to quantify the total uncertainty under a varying climate?

These are the main questions studied herein. 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 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 demonstration 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.

The classical statistical estimators (e.g. for the mean, variance, etc.) 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. A more consistent representation, which relaxes, to some degree, the assumption of stable climate, is studied in the next section with the help of real-world examples. In the subsequent sections, typical statistical tasks such as estimation, prediction and hypothesis testing are revisited, both theoretically and by means of case studies.

STOCHASTIC REPRESENTATION OF HYDROMETEOROLOGICAL PROCESSES

Varying climate and the Hurst phenomenon

To establish a pragmatic representation of hydrometeorological processes, it is very useful to examine observed long time series. Three example time series have been chosen as case studies, which are depicted in Figs 1–3. The first example (Fig. 1(a)) is a long time series (992 years) representing the Northern Hemisphere temperature anomalies in °C *vs* 1961–1990 mean (Jones *et al*., 1998a,b). This series was constructed

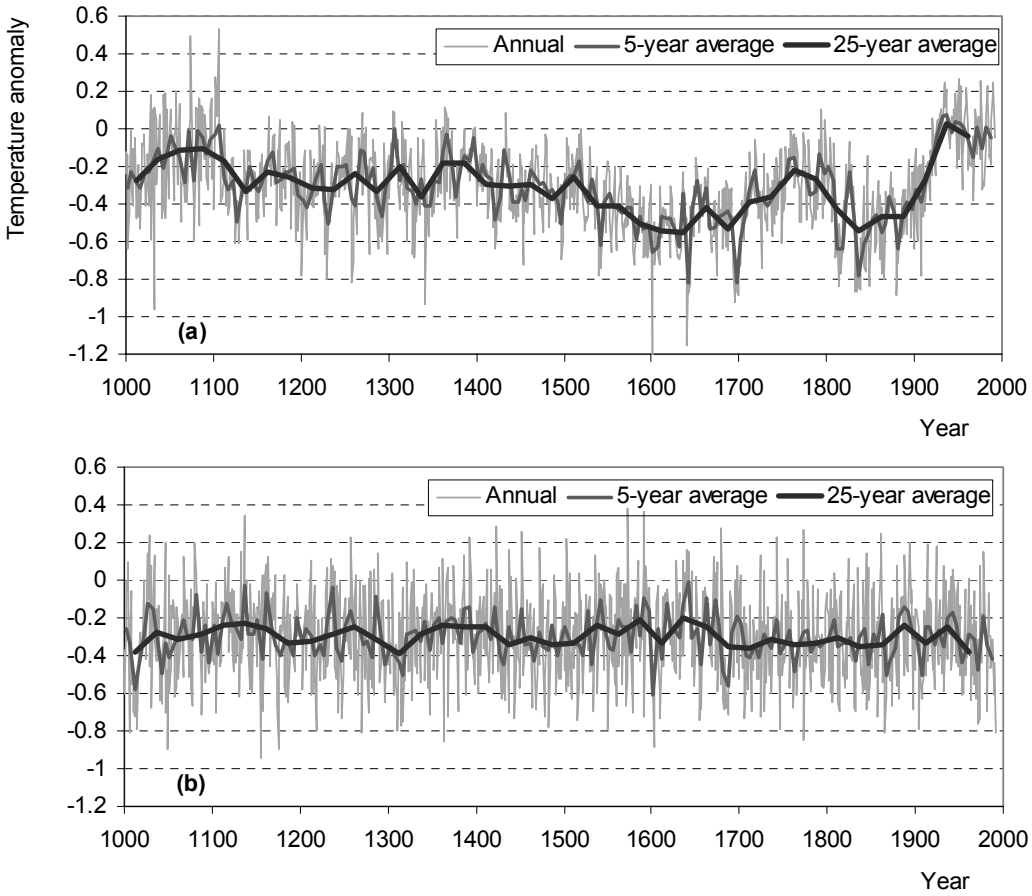

Fig. 1 (a) Plot of the Northern Hemisphere temperature anomalies, reconstructed by *Jones et al.* (1998b) using proxy data; and (b) for comparison, a series of white noise with mean and standard deviation equal to those of the original series is also plotted.

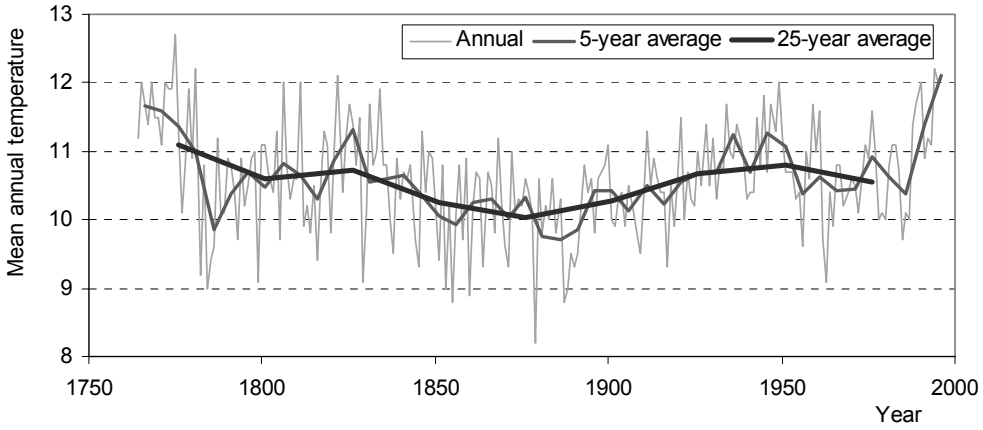

Fig. 2 Plot of the time series of mean annual temperature at Paris/Le Bourget.

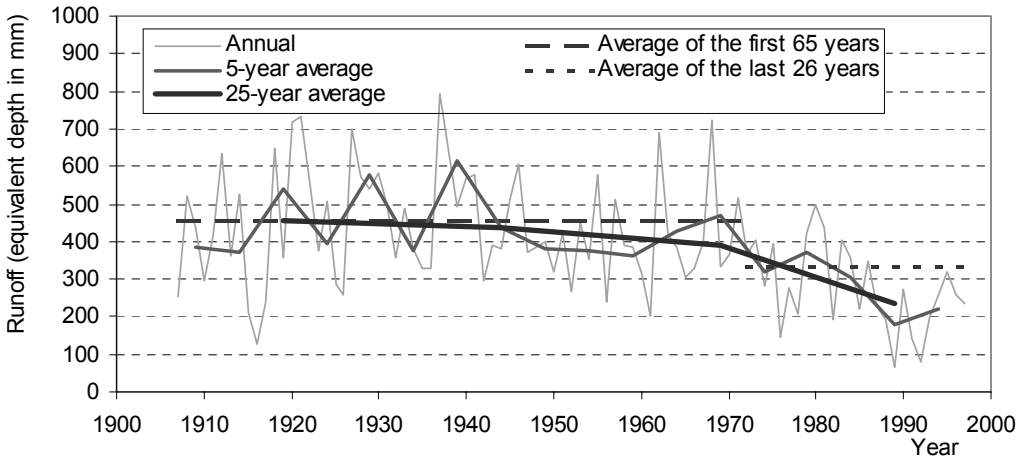

Fig. 3 Plot of the time series of the equivalent runoff depth of the Boeotikos Kephisos River basin, Greece.

using temperature sensitive palaeoclimatic multi-proxy data from 10 sites worldwide that include tree rings, ice cores, corals, and historical documents. The second example (Fig. 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 (from ftp.cru.uea.ac.uk). The third example (Fig. 3) is the longest available streamflow record in Greece, the Boeotikos Kephisos River runoff. The river is located to the north of Athens and the time series length is 91 years (hydrological years 1907/08–1997/98).

A common characteristic in all three examples is that a local multi-year 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 (Fig. 1(a)), during the 16th century, there is a falling trend of the local average, which is inverted during *c.* 1650–1750, becomes again falling during *c.* 1750–1850, and becomes rising thereafter. 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 Fig. 1(b). The difference of the original series from white noise is clearly visible; the variability of the latter in the aggregation level of 25 years is much lower than that of the former.

Similar behaviour has been reported in several long time series. For example Beran (1994, p. 41), examined a number of geophysical and other long time series and observed that, at short time periods, there seem to be local trends, but, looking at the whole series, these become cycles of (almost) all frequencies, superimposed and in random sequence.

The falling and rising local trends can be regarded as climate changes or variations, considered by many as deterministic components in climatic time series. In all cases, however, these changes are irregular and, in the absence of an accurate deterministic model that could explain them and predict their future, are better modelled as stochastic fluctuations on many time scales. Equivalently, these fluctuations can be regarded as a manifestation of the Hurst phenomenon discovered by Hurst (1951). This phenomenon is quantified through the so-called Hurst exponent (or coefficient), H, the mathematics of which is discussed later. In a white noise series, such as the one plotted in Fig. 1(b), H takes the value 0.5 whereas in real-world time series H is usually greater. The coefficient H provides a direct means to demonstrate the close relationship of large-scale variability with the Hurst phenomenon: if H is calculated using such a time series the result will be higher than 0.5 but, if the same calculation is repeated for the "detrended" series (i.e. the series after "subtracting" the *a posteriori* identified trends), H will be around 0.5.

In a recent paper (Koutsoyiannis, 2002), this relationship was more thoroughly explored and it was demonstrated that the multiple-scale variability of a time series can explain the Hurst phenomenon. This explanation is consistent with observations by Mesa & Poveda (1993), who regarded the Hurst phenomenon as "probably the result of a mixture of scales" and by Vanmarcke (1983, p. 225), who studied a composite random process with two different scales of fluctuation.

Having hypothesized that the Hurst phenomenon is a manifestation of irregular climatic fluctuations on several scales, it can be conjectured that identifying the Hurst phenomenon on a specific series provides some indication of climatic fluctuations. A number of studies have identified the Hurst phenomenon in several environmental quantities such as (to mention a few of the more recent) wind power (Haslet & Raftery, 1989); global mean temperatures (Bloomfield, 1992); flows of the Nile (Eltahir, 1996; Koutsoyiannis, 2002); flows of the River Warta, Poland (Radziejewski & Kundzewicz, 1997); inflows of Lake Maggiore, Italy (Montanari *et al.*, 1997); indexes of North Atlantic Oscillation (Stephenson *et al.*, 2000); and tree-ring widths, which are indicators of past climate (Koutsoyiannis, 2002).

In conclusion, a stochastic representation of hydrometeorological time series that respects the Hurst phenomenon is conjectured to be consistent with the varying climate hypothesis.

Basic assumptions and notation

Let X_i denote a hydrometeorological process with $i = 1, 2, \ldots$, denoting discrete time. In the context of this paper it is assumed that X_i is not periodic, which means that time

scale is at least annual. Rather, it is assumed that the process is stationary, a property that does not hinder to exhibit multiple-scale variability. Further, let its mean be denoted as $\mu := E[X_i]$, its autocovariance $\gamma_j := \mathrm{cov}[X_i, X_{i+j}]$, its autocorrelation $\rho_j := \mathrm{corr}[X_i, X_{i+j}] = \gamma_j/\gamma_0$ ($j = 0, \pm1, \pm2, \ldots$), and its standard deviation $\sigma := \sqrt{\gamma_0}$.

Let k be a positive integer that represents a time scale greater than the basic time scale of the process, X_i. The aggregated stochastic process on that time scale is denoted as:

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

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

$$E[Z_i^{(k)}] = k\mu \tag{2}$$

whilst the variance and autocovariance (or autocorrelation) depend on the specific structure of γ_j (or ρ_j). The hypothesis is employed that hydrometeorological processes exhibit scale invariant properties at any scale greater than annual, i.e.:

$$(Z_i^{(k)} - k\mu) \overset{\mathrm{d}}{=} \left(\frac{k}{l}\right)^H (Z_j^{(l)} - l\mu) \tag{3}$$

where the symbol $\overset{\mathrm{d}}{=}$ stands for equality in (finite dimensional joint) distribution. Equation (3) is valid for any integer i and j (that is, the process is stationary) and any time scales 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). It is easily shown (e.g. Bras & Rodriguez-Iturbe, 1985, p. 221) that the process defined by equation (3) reproduces the Hurst phenomenon. In the scope of this paper, the standard name "fractional Gaussian noise" is avoided for several reasons: the first term, fractional, is not easily understandable, unless combined with fractals, which is not necessary when dealing with hydrological statistics. The second term, Gaussian, may be not appropriate for several hydrological processes that are asymmetric (non-Gaussian). There is no need to restrict the analysis to Gaussian processes, and techniques to produce synthetic time series respecting equation (3) that are not Gaussian have been studied in Koutsoyiannis (2000, 2002.) The third term, noise, usually describes a random and unstructured process, which is not the case in hydrological processes that are structured. An alternative name for the process (equation (3)) is "stationary increments of self-similar process" (Beran, 1994, p. 50), which is more general but rather complicated. Here, it was preferred to use the name simple scaling stochastic process or simple scaling signal (SSS).

As a consequence of equation (3), for $i = j = l = 1$ one obtains that the variance of the aggregated process is:

$$\gamma_0^{(k)} := \mathrm{var}[Z_i^{(k)}] = k^{2H}\gamma_0 \tag{4}$$

Thus, the standard deviation $\sigma^{(k)} := (\gamma_0^{(k)})^{1/2}$ is a power law of the scale or level of aggregation k with exponent H. The autocorrelation function, for any aggregated time

scale k, is independent of k, and given by:

$$\rho_j^{(k)} = \rho_j = (1/2)\,[(j+1)^{2H} + (j-1)^{2H}] - j^{2H} \approx H(2H-1)j^{2H-2} \qquad j > 0 \tag{5}$$

STATISTICAL ESTIMATION AND PREDICTION UNDER THE SSS REPRESENTATION

Estimations for the most common statistics that are used in hydrological estimation, prediction and testing are studied here under the hypothesis that the process of interest is SSS. It is assumed that the sample is a time series of length n whose items correspond to consequent time instances, i.e. $X_1, \ldots, X_n$.

Estimation of mean

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

$$\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i \tag{6}$$

As can be directly verified, $\overline{X}$ is an unbiased estimator, regardless of the type of the process X_i, i.e. $E[\overline{X}] = \mu$. Moreover, it is very close to the best linear unbiased estimator of the process mean for SSS (Adenstedt, 1974; Beran, 1994, p. 150). In classical statistics, its variance is:

$$\mathrm{var}[\overline{X}] = \frac{\sigma^2}{n} \tag{7}$$

which, however, is not valid in SSS. Instead, observing that $\overline{X} = Z_1^{(n)}/n$ and using equation (4), one obtains:

$$\mathrm{var}[\overline{X}] = \frac{\sigma^2}{n^{2-2H}} \tag{8}$$

This is a known result for SSS (Adenstedt, 1974; Beran, 1994, p. 54). It is reminded that the square root of $\mathrm{var}[\overline{X}]$ is the standard error in estimating the true mean from the observed time series. For $H = 0.5$ both equations (7) and (8) result in the same standard error, which is inversely proportional to the square root of the length of the time series. However, for large values of H, the difference between equations (7) and (8) becomes very significant. For example, in a time series of $n = 100$ years observations, according to the classical statistics, the standard estimation error is $\sigma/10$. However, for $H = 0.8$ the correct standard error, as given by equation (8), is $\sigma/2.5$, i.e. four times larger. To have an estimation error equal to $\sigma/10$, 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 will be seen below.

Estimation of variance and standard deviation for known Hurst coefficient

The classical variance estimator

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \tag{9}$$

is no longer an unbiased estimator for SSS. It has been shown (Beran, 1994, p. 156) that a consistent SSS estimator, which becomes unbiased for known H, is:

$$\widetilde{S}^2 := \frac{n-1}{n-n^{2H-1}}S^2 = \frac{1}{n-n^{2H-1}}\sum_{i=1}^{n}(X_i - \bar{X})^2 \tag{10}$$

When $H = 0.5$, equation (10) becomes identical to equation (9). This expression has some similarity with earlier expressions that corrected the variance estimator in terms of the lag-1 autocorrelation (Salas, 1993).

The statistic $\widetilde{S}$ (the square root of $\widetilde{S}^2$) can then be considered as an estimator of the standard deviation σ. However, since the square root (as any nonlinear transformation) does not preserve unbiasedness, $\widetilde{S}$ is biased. A more consistent estimator (approximately unbiased for known H and for normal distribution of X_i) is:

$$\widetilde{\widetilde{S}} := \sqrt{\frac{n-1/2}{n-n^{2H-1}}}\,S = \sqrt{\frac{n-1/2}{(n-1)(n-n^{2H-1})}}\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{11}$$

This was based on a systematic Monte Carlo study. It can be easily verified that, when $H = 0.5$, $\widetilde{\widetilde{S}} = [(n - ½)/(n - 1)]^{0.5}S$. This is a very good approximation of the true unbiased estimator, which for independent Gaussian X_i is:

$$\widetilde{\widetilde{S}} = \{[(n-1)/2]^{0.5}\,\Gamma[(n-1)/2]/\Gamma(n/2)\}S$$

The variance of S for a normal distribution of X_i in the classical statistics is:

$$\mathrm{var}[S] \approx \frac{\sigma^2}{2(n-c)} \tag{12}$$

where c is taken typically as 0 or 1 but, more accurately, can be taken as 0.75 (the exact variance is $\{1 - [2/(n-1)]\Gamma^2(n/2)/\Gamma^2[(n-1)/2]\}\,\sigma^2$. In the SSS case, with the same Monte Carlo study, it was found that:

$$\mathrm{var}[\widetilde{\widetilde{S}}] \approx \frac{(0.1n + 0.8)^{\lambda(H)}\,\sigma^2}{2(n-1)} \quad \text{with } \lambda(H) := 0.088(4H^2 - 1)^2 \tag{13}$$

It can be verified that equation (13) results in higher variance than that given by equation (12) (apart from the case $H = 0.5$, when the two equations become almost identical). It is emphasized that both equations (11) and (13) refer to normal distribution of X_i. Preliminary investigation showed that equation (11) can be also used for gamma distributed X_i but the variability in this case is higher than the one given by equation (13).

A demonstration of the consequences of using the inappropriate classical estimators of variance and standard deviation is given in Koutsoyiannis (2003). In brief, equation (9) results in underestimation of variance, which for small time scales is not serious, but increases for larger scales. The increasingly underestimated standard deviation with the increase of scale k, when the classical estimator is used, has another important consequence: it obscures the presence of the Hurst phenomenon for small samples. Specifically, in a logarithmic plot of standard deviation *vs* scale, the slope of the curve of classical estimate is not constant, as implied by the scaling law (equation (4)), 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 to the wrong conclusion that for large scales this slope tends to 0.5, a value indicating absence of the Hurst phenomenon.

Simultaneous estimation of variance and Hurst coefficient

When the Hurst exponent is unknown, which is the usual case when dealing with an observed time series, equations (10)–(11) cannot be applied as they contain the unknown H. Therefore, a new algorithm, consistent with the SSS statistics, is proposed. This algorithm is based on classical sample estimates $s^{(k)}$ of standard deviations $\sigma^{(k)}$ for time scales 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 equations (11) and (4) and assuming $E[\tilde{\tilde{S}}] = \sigma$, one obtains:

$$s^{(k)} \approx c_k(H) k^H \sigma \qquad \text{with } c_k(H) := \sqrt{\frac{n/k - (n/k)^{2H-1}}{n/k - 1/2}} \tag{14}$$

Ignoring $c_k(H)$ in equation (14), one can estimate H and σ by linear regression of $\ln s^{(k)}$ on $\ln k$. A graphical depiction of this regression is always very important, as it provides visual evidence of the appropriateness of the SSS model for the time series of interest. This regression is a known algorithm (e.g. Montanari *et al.*, 1997; Koutsoyiannis, 2002), which, however, introduces negative bias for both H and σ when H is high. To reduce bias, in a second trial, the estimate of H from the regression can be used to determine $c_k(H)$ and then perform another regression of $\ln s^{(k)}$ on $\ln[c_k(H)k]$. This procedure can be continued until convergence.

A more systematic approach can be formulated in terms of minimizing a fitting error, i.e.:

$$e^2(\sigma, H) := \sum_{k=1}^{k'} \frac{[\ln\sigma^{(k)} - \ln s^{(k)}]^2}{k^p} = \sum_{k=1}^{k'} \frac{[\ln\sigma + H\ln k + \ln c_k(H) - \ln s^{(k)}]^2}{k^p} \tag{15}$$

where a weight equal to $1/k^p$ was assigned to the partial error of each scale k. For $p = 0$ the weights are equal, whereas for $p = 1, 2, \ldots$, decreasing weights are assigned to increasing scales; this is reasonable because, at larger scales, the sample size is smaller and thus uncertainty larger. Using Monte Carlo experiments it was found that, although differences in estimates caused by different values of p in the range 0–2 are

not so important, $p = 2$ results in slightly more efficient estimates (i.e. with smaller variation) and thus is preferable. From equations (14) and (15) it can be observed that, when σ tends to zero or infinity, or when H tends to 1, $e^2(\sigma, H)$ tends to infinity. This guarantees that a global minimum exits, in which $\sigma > 0$ and $H < 1$. However, an analytical procedure to locate it is very difficult to set up. Therefore, minimization of $e^2(\sigma, H)$ should be done numerically and several numerical procedures can be devised for this purpose. A detailed iterative procedure is given in Koutsoyiannis (2003).

To demonstrate the performance of the algorithm, a Monte Carlo experiment has been carried out using an ensemble of 100 samples with $n = 50$ and 100, generated for $H = 0.80$, $\mu = 2$ and $\sigma = 0.50$. For comparison, similar series of white noise ($H = 0.50$) were also used. The resulting values of H and σ using the above algorithm are shown in Fig. 4 by means of box plots. The figure shows a good performance of the algorithm for estimation of both H and σ. For comparison, the algorithm based on the regression method mentioned above, which essentially is based on the classical estimator of σ, has been also applied. As shown in Fig. 4, this 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, Fig. 4 shows that this algorithm is inappropriate. It exhibits a negative bias and, 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 as in H smaller than 0.5, which is physically inconsistent (although mathematically acceptable). In contrast, the proposed algorithm, as discussed earlier, results in $H < 1$; in the simulation, it never resulted in H smaller than 0.5 when the true H is 0.8.

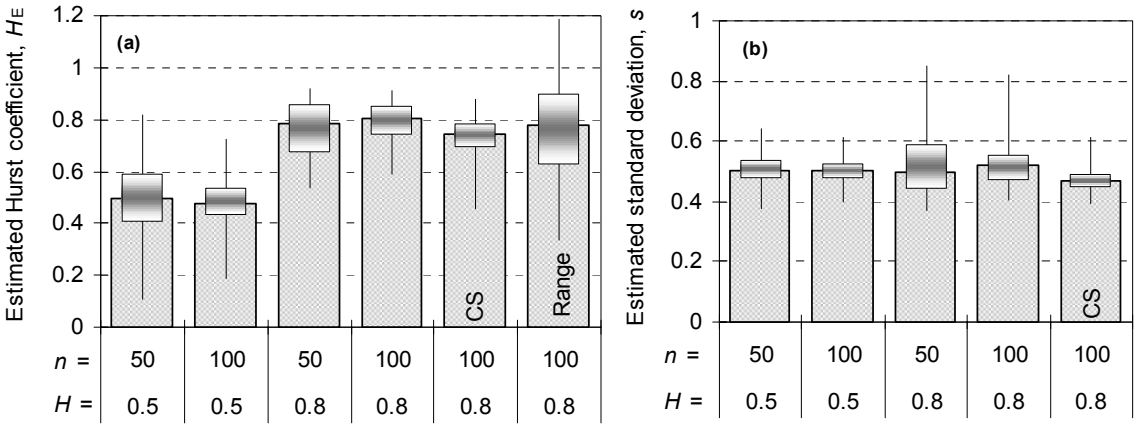

Fig. 4 Box plots of the estimated (a) Hurst coefficients and (b) standard deviations from ensembles of synthetic series with theoretical $H = 0.5$ or 0.8 and $\sigma = 0.5$. Bars correspond to the median of 100 estimations (from 100 synthetic series), upper and lower edges of boxes correspond to the 75 and 25% quantiles, respectively, and whiskers correspond to the maximum and minimum estimated values. The first four boxes correspond to the proposed method, the fifth box (annotated as CS) corresponds to estimates using classical statistics, and in panel (a) the sixth box (annotated as Range) corresponds to the estimation using the original Hurst's algorithm based on the rescaled range.

Estimation of distribution quantiles

The problem of estimation of quantiles of hydrological variables, for certain values of probability of nonexceedence, is central to hydrological statistics. Clearly, the above results lead to dramatic differences of these quantiles in comparison with their classical estimates.

To demonstrate this, the case of normally distributed X_i is considered, where the classical estimator of the u quantile (the value of the variable for probability of nonexceedence u) is:

$$\hat{X}_u = \overline{X} + \zeta_u S \tag{16}$$

where ζ_u is the u quantile of the standard normal distribution. Its classical confidence limits for confidence coefficient, γ, are given by (e.g. Stedinger *et al.*, 1993):

$$\hat{x}_{u_{1,2}} = \hat{x}_u \pm \zeta_{(1+\gamma/2)}\varepsilon_u \qquad \text{with } \varepsilon_u = \frac{s}{\sqrt{n}}\sqrt{1+\frac{\zeta_u^2}{2}} \tag{17}$$

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

$$\hat{Z}_u^{(k)} = k\overline{X} + \zeta_u k^H \tilde{\tilde{S}} \tag{18}$$

The confidence limits can be estimated using equations (8) and (13) and also assuming that, for normally distributed X_i, $\overline{X}$ and $\tilde{\tilde{S}}$ are independent (as in the classical case), a fact verified by Monte Carlo simulations. After algebraic manipulations, these confidence limits become:

$$\hat{z}_{u_{1,2}}^{(k)} = \hat{z}_u^{(k)} \pm \zeta_{(1+\gamma/2)}\varepsilon_u^{(k)} \qquad \text{with } \varepsilon_u^{(k)} = k\frac{\tilde{\tilde{s}}}{n^{1-H}}\sqrt{1+\frac{\zeta_u^2\,(0.1n+0.8)^{\lambda(H)}}{2(k/n)^{2-2H}(n-1)}} \tag{19}$$

To demonstrate the differences between the classical and SSS estimators, the estimates of quantiles from the first 10 synthetic samples of the ensemble already described in the previous subsection for n = 100 have been plotted (on Gauss probability paper) in Fig. 5 for both cases, classical and SSS (Fig. 5(a) and (b), respectively). This figure is for the basic time scale (k = 1). The theoretical distribution function is also plotted along with 95% confidence limits around it that correspond to the true mean and standard deviation equal to the largest of the 10 sample standard deviations (0.59 and 0.77 for the classical and the SSS case, respectively). It is observed in Fig. 5 that the confidence interval determined by the SSS estimator is 3–5 times wider than that of the classical estimator. Even though the largest of the 10 sample standard deviations was used to determine the confidence limits, more than half of the 10 empirical quantile curves lie outside of the 95% classical confidence limits for the probability domain shown in the graph. In contrast, all curves lie within the 95% SSS confidence limits. This demonstrates the inappropriateness of classical estimators and the appropriateness of the SSS ones.

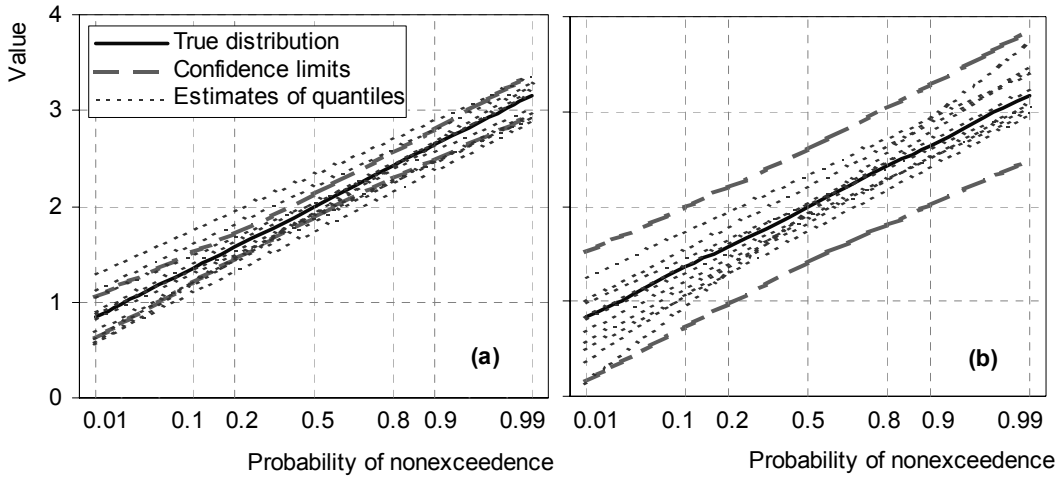

Fig. 5 Estimates of quantiles at the basic time scale (k = 1) from 10 synthetic time series with length n = 100 generated with theoretical parameters H = 0.8, μ = 2 and σ = 0.5, in comparison with the true distribution and 95% confidence limits: (a) using classical statistics; (b) using SSS statistics.

When both the standard deviation and the Hurst exponent are unknown, the confidence intervals are even wider. However, their theoretical determination is very difficult and therefore Monte Carlo methods should be the appropriate choice.

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, it can be shown (see Koutsoyiannis, 2003) that: (a) the typical cross-covariance estimator is biased; (b) a consistent estimator can be formed with an expression analogous to equation (10); and (c) the classical estimator of the cross-correlation coefficient remains valid also for SSS. These results have been verified by Monte Carlo experiments with bivariate time series generated using the multivariate method described by Koutsoyiannis (2000).

Estimation of autocovariances and autocorrelations

It is known that for series with nonzero autocorrelation, the typical estimator of auto-covariance is biased downward (Salas, 1993). It can be shown that the typical estimator of the lag l autocovariance:

$$G_l := \frac{1}{n}\sum_{i=1}^{n-l}(X_i - \bar{X})(X_{i+l} - \bar{X}) \tag{20}$$

is biased for SSS processes and an approximately unbiased estimator of γ_l is:

$$\widetilde{G}_l := G_l + \frac{1}{n^{2-2H}}\widetilde{S}^2 = G_l + \frac{n-1}{n^{3-2H}-n}S^2 \tag{21}$$

(a proof can be found in Koutsoyiannis, 2003). Consequently, a consistent estimator of

the autocorrelation coefficient ρ_l will be:

$$\widetilde{R}_l := \frac{\widetilde{G}_l}{\widetilde{S}^2} = R_l\left(1 - \frac{1}{n^{2-2H}}\right) + \frac{1}{n^{2-2H}} \tag{22}$$

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

$$R_l := \frac{n}{n-1}\frac{G_l}{S^2} \tag{23}$$

Both equations (21) and (22) are in agreement with the asymptotic results that were determined by Hosking (1996) for a normally distributed SSS and for $n \to \infty$.

Clearly, the classical estimate of autocorrelation is highly biased for large H. This is demonstrated in Fig. 6 by means of the Monte Carlo experiment already discussed

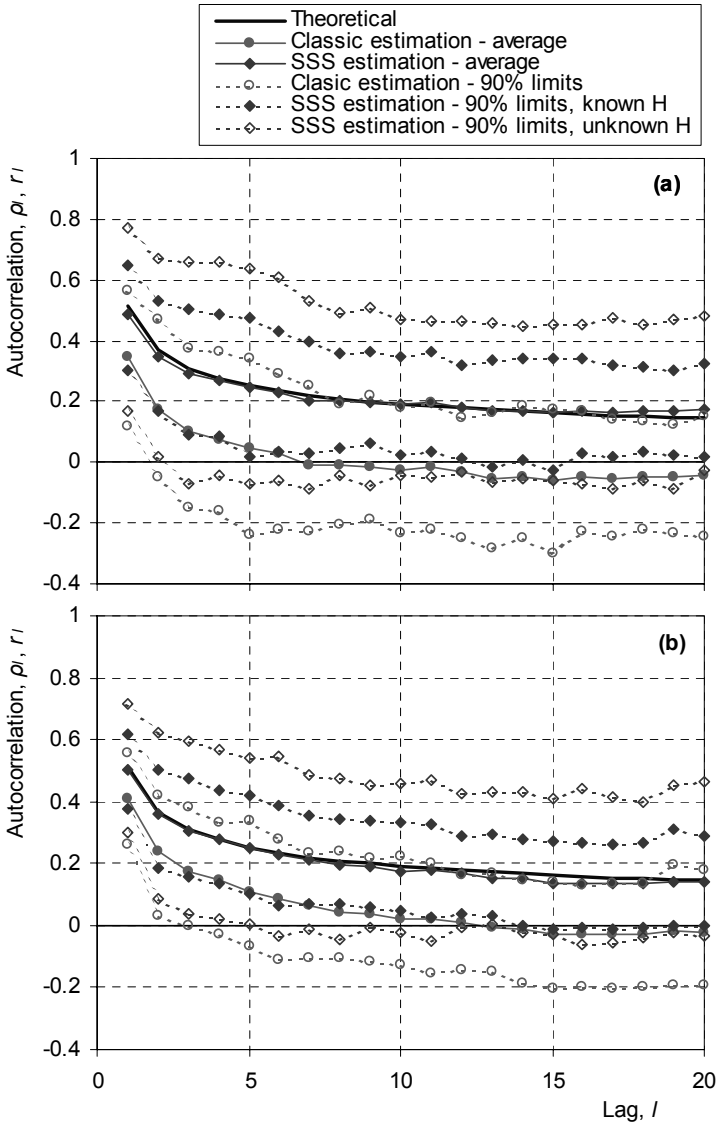

Fig. 6 Comparison of theoretical and empirical autocorrelation functions at the basic scale (k = 1) for a Monte Carlo experiment with theoretical H = 0.8 and σ = 0.5 for (a) n = 50; and (b) n = 100.

earlier. Not only are the classical estimates of autocorrelation coefficient significantly lower than the true values, but they also vanish 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. In contrast, the SSS estimator captures very well the long-term persistence of the process and agrees well with the theoretical autocorrelation function. Apart from the point estimates of autocorrelations, 90% confidence limits, estimated from simulation, have been also plotted in Fig. 6. It can be observed that, when H is regarded as known (H = 0.80), the SSS confidence zone is as wide as the classical one but significantly shifted upward. However, when H is regarded as unknown (with different value in each sample estimated using the proposed algorithm), the SSS confidence zone becomes significantly wider. For instance, for lag 5 and for n = 100, the 90% confidence zone extends between 0 and 0.55. This indicates the vastly high uncertainty in estimating the autocorrelograph, even for a sample size as large as 100.

CASE STUDIES

The three real-world time series examined above are discussed further here. In Fig. 7, the standard deviation of Jones's (1998b) proxy time series of temperature anomalies has been plotted *vs* time scale. Clearly, the standard deviation of the aggregated process is a power function of time scale and this is apparent, even using the classical estimator of standard deviation. The exclusion of the data of the last century, which can be suspect for reasons of anthropogenic influence, does not alter the shape of the curve of standard deviation *vs* slope, as also shown in Fig. 7. The slope of this curve in the log-log plot is much greater than 0.5. The estimate of Hurst coefficient is 0.86 using the regression method and 0.88 using the proposed method.

In Fig. 8(a) the lag-1 and lag-2 autocorrelation coefficients have been plotted *vs* scale. The SSS model implies that these autocorrelations are independent of scale (equation (5)). The classical empirical estimates of autocorrelation verify this

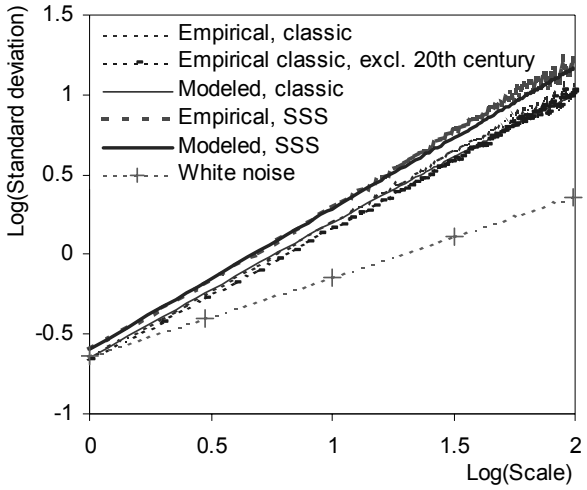

Fig. 7 Standard deviation of the aggregated processes *vs* time scale (logarithmic plot) for Jones's time series of the Northern Hemisphere temperature anomalies. For comparison, the theoretical curve of the white noise model is also plotted.

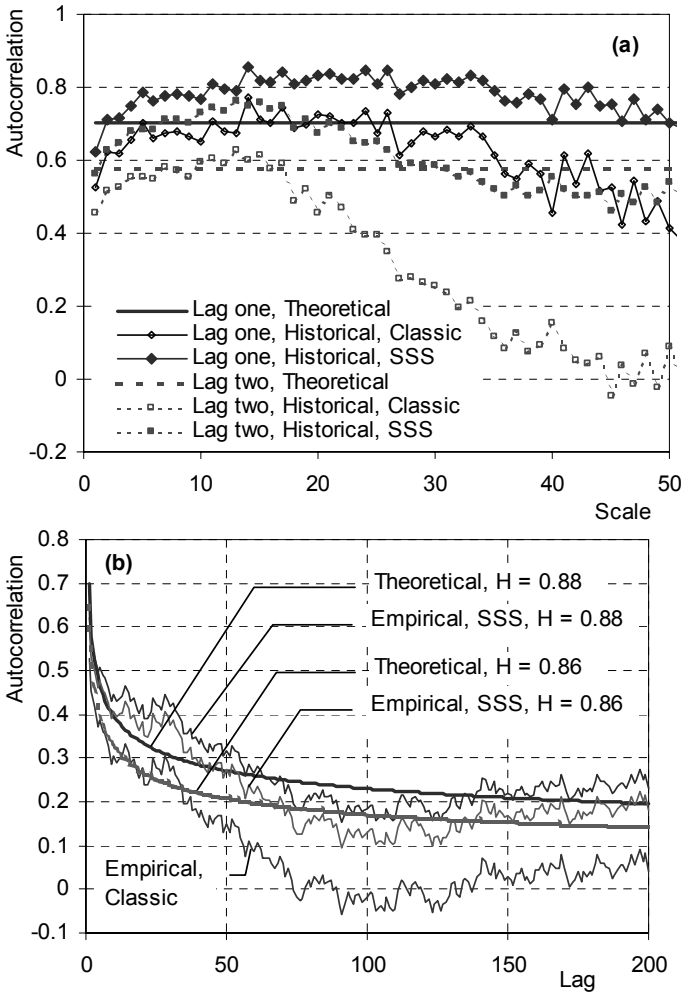

Fig. 8 Autocorrelation coefficients of Jones's time series of the Northern Hemisphere temperature anomalies: (a) lag-1 and lag-2 autocorrelations of the aggregated process *vs* time scale, k; and (b) autocorrelation *vs* lag for the basic time scale, $k = 1$.

theoretical expectation, but not fully for very large scales (>20), where autocorrelation decreases somehow. However, the SSS estimates agree well with the model, as they are almost constant even for scales as large as 50 years. In Fig. 8(b) the autocorrelation function for the basic scale ($k = 1$) has been plotted *vs* lag. Due to the large length of the time series, the long-term persistence of the time series is obvious here even when the classical autocorrelation estimator is used. However, in this case, the empirical autocorrelation departs from theoretical ones for lags >40, even when the value $H = 0.86$ is used. When the SSS estimate is used with the value $H = 0.88$, the model fits satisfactorily to the empirical autocorrelation function for lags up to 200.

In Fig. 9, the point estimates and the 99% confidence limits of the quantiles of the temperature anomalies have been plotted for probability of nonexceedence, u, ranging from 1 to 99%, assuming a normal distribution as verified from the time series. This is done for two time scales, the basic one ($k = 1$) that represents the annual variation of temperature anomaly, and the 30-year time scale, which typically is assumed to be sufficient to smooth out the annual variations and provide values representative of the climate. (For the latter, the averaged rather than aggregated time series, i.e. $z_i^{(30)}/30$,

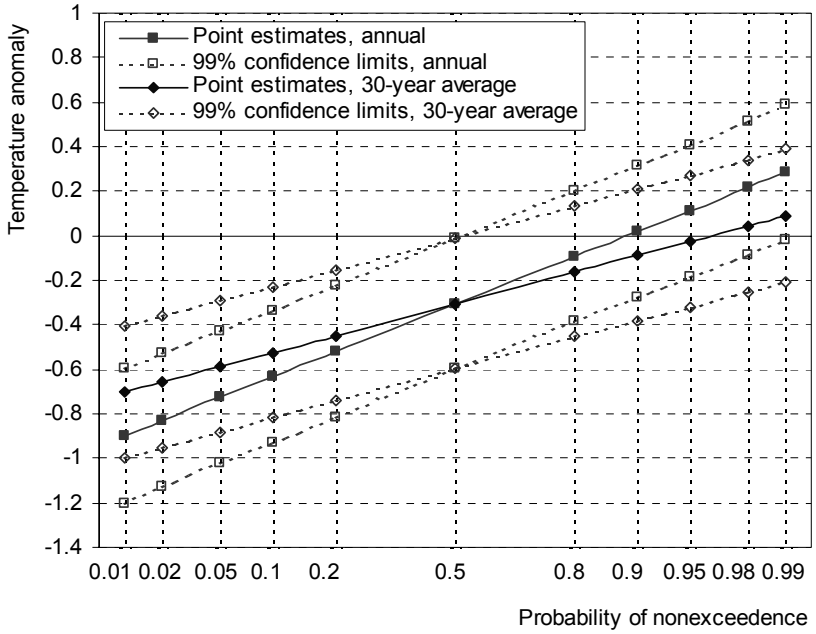

Fig. 9 Point estimates of quantiles at the basic time scale (annual values, $k = 1$) and the 30-year time scale (30-year averages, $k = 30$), and 99% confidence limits thereof for Jones's time series of the Northern Hemisphere temperature anomalies.

has been used.) It is observed in Fig. 9 that the variation of the 30-year average is only slightly lower than that of the annual values. For example, the point estimate of the 99% quantile of annual temperature anomaly is 0.6°C above average, whereas that of the 30-year average is 0.4°C above average (which over all 992 years is not zero, but rather –0.30°C, as in this time series the temperature anomalies are expressed as differences from the 1961–1990 mean). If one considers the upper 99% confidence limits of quantiles, these values become about 0.9 and 0.7°C above average, respectively. These figures indicate that an increase of the 30-year average temperature by 0.5°C, does not provide strong statistical evidence of an unusual change of climate. (It is noted that the observed temperature increase since 1850 is around 0.5°C.)

The results related to this example contain a high degree of uncertainty due to the proxy character of the time series. However, this problem does not emerge in the next example, the Paris temperature time series. In Fig. 10(a), the standard deviation of Paris temperature has been plotted *vs* time scale. Again, the standard deviation of the aggregated process is a power function of time scale and this is apparent even using the classical estimator of standard deviation. The Hurst coefficient, estimated using the proposed method, is 0.79. The autocorrelation coefficients also verify the presence of long-term persistence (a plot for this time series similar to that of Fig. 8 can be found in Koutsoyiannis, 2003). Otherwise, as shown in Fig. 2, there is no remarkable pattern in this time series that would require further statistical analysis.

An interesting pattern exists in the third example time series, the Boeotikos Kephisos runoff, shown in Fig. 3. This is the falling trend since 1920, lasting for 78 out of a total of 91 years. The characteristic plot of standard deviation *vs* time scale, shown in Fig. 10(b), again verifies the presence of long-term persistence. As in the previous examples, the relationship of standard deviation to scale is a power law; the Hurst exponent is 0.79.

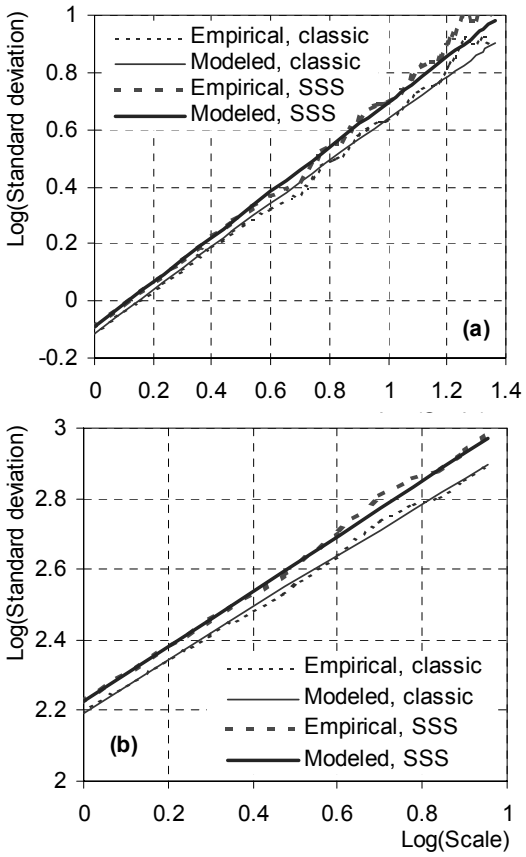

Fig. 10 Standard deviation of the aggregated processes *vs* time scale (logarithmic plot) for the time series of (a) mean annual temperature at Paris/Le Bourget; (b) runoff time series of the Boeotikos Kephisos River basin.

Typically, in hydrological statistics, a trend is detected using the Kendall's τ statistic (e.g. Kottegoda, 1980, p. 32) defined as:

$$\tau := \frac{4p}{n(n-1)} - 1 \tag{24}$$

where p is the number of 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 this example, 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 exists. However, this is not correct, because the time series is not a series of independent values but it exhibits SSS properties. To find the value of the standard deviation of τ, stochastic simulation can be used. Thus, an ensemble of 100 time series with $n = 78$ and $H = 0.78$ was generated, from which it was found that the standard deviation of τ is 0.173 (more than twice greater than 0.077) and the attained significance level of Kendall's test now becomes 0.01. Still, this is not absolutely correct because it does not correspond exactly to the formulation of the null and alternative hypotheses. It is recalled that, to formulate the hypotheses to be tested, an

exploratory data analysis of the complete 91-year series was carried out and this trend was located in 78 of the 91 years. Further, 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, a different stochastic simulation that is consistent with the procedure of formulating the hypotheses was performed. Specifically, an ensemble of 100 time series was generated, each with $n = 91$ and $H = 0.78$, and in each of these series that 78-year period which gave the maximum value of τ (in absolute value) was located. Now the standard deviation of τ over the 100 series is 0.252 (more than three times greater than 0.077) and the attained significance level of 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 Fig. 3, can be located between 1971 and 1972. The 65-year period before the jump and the 26-year period after the jump have averages of 439.3 and 276.6 mm, respectively, whereas the entire 91-year average is 392.8 mm. In classical 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}. However, this typical test is incorrect because it was based on the classical statistics. To perform a more accurate test, stochastic simulation was performed. In each of the 100 generated 91-year long synthetic series already discussed above, the 26-year period with the minimum average was located and the difference from the average of the entire 91-year series was taken. Then, the probability that this difference exceeds 392.8 – 276.6 = 116.2 mm was determined, 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 Boeotikos Kephisos 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 classical statistics.

SUMMARY, CONCLUSIONS AND DISCUSSION

According to the experts of climate modelling, "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 be 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".

Therefore, probability-based methods may be good alternatives to quantify uncertainty, even under a varying climate. However, hydrological 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 changed irregularly on all time scales throughout the history of the Earth, as witnessed by 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 could be processed using classical statistics. This would be an efficient approach if a deterministic model existed, which could explain these components and also predict their future. However, this is hardly the case, as most typically the trends or shifts are identified only *a posteriori* and expressed mathematically by equations lacking physical meaning (e.g. using linear regression) and thus applicable only to the relevant parts of the time series and not their future evolution. A more consistent alternative is to approach this behaviour in a stochastic manner. A stochastic basis for dealing with these shifts and trends is offered by simple scaling processes that are consistent with the assumption of hydroclimatic fluctuations on multiple time scales, a behaviour that is none other than the Hurst phenomenon. The Hurst coefficient is the exponent of the power-law relationship between the aggregated standard deviation at any time scale and the scale length.

When the simple scaling hypothesis is employed and the typical statistical descriptors used in hydrological statistics are revisited under this hypothesis, it is found that:

- the classical sample average remains an unbiased estimator of the true mean, but its variance, which expresses the uncertainty of its estimation, is dramatically higher than the value given by the classical statistics;
- the classical estimators of variance and standard deviation are no more unbiased and their variance is larger than that of the classical statistics;
- the quantiles of a given distribution function differ from those of the classical statistics and their confidence intervals are radically wider than those implied by the classical statistics;
- the classical estimators of cross-correlations between two variables remain almost unbiased, but those of autocorrelations are highly biased downward; the confidence intervals of the latter are much wider than those implied by the classical statistics.

For all the above statistical descriptors, generalized estimators exist 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 classical statistics. The classical 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, simultaneously with standard deviation, in an approximately unbiased manner.

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 are consistent with the scaling hypothesis (Hurst phenomenon). In addition, it is shown that several patterns within these time series would be regarded as evident trends or shifts if classical statistical tests were used, but using modified tests, based on the scaling hypothesis, it turns out that it is a regular behaviour.

Apparently, the consistency of geophysical time series with the scaling hypothesis is not exhausted for the three time series analysed herein. In several studies, a large number of long geophysical time series have been found to exhibit the Hurst phenomenon. The analyses herein show that in time series with short length, the classical statistics can hide the scaling behaviour. This concerns the Hurst exponent and the

autocorrelation function (Fig. 6), whose classical estimate hides a fat tail. Therefore, it can be the case that short time series, classified as random noise without scaling behaviour, in fact exhibit the Hurst phenomenon.

In conclusion, the analyses in this study act as a warning that the classical hydrological statistics describes only a portion of the natural uncertainty of hydroclimatic processes, because it is based on the implicit assumption of a stable climate. In addition, its use may characterize a regular behaviour of hydroclimatic processes as an unusual phenomenon. Furthermore, the analyses show that it is feasible to adapt the classical hydrological statistics so as to quantify (but not to reduce) the total uncertainty under a varying climate. In some cases, the adapted estimators are based on simplifying assumptions, such as the assumption of normal distribution of the process of interest. It may be difficult to derive closed analytical estimators appropriate for any type of distribution and for every statistic involved in hydrological estimation and testing. However, stochastic simulation provides an easy and directly applicable means for problems that cannot be solved analytically, as demonstrated in the examples of hypothesis testing presented here. Obviously, further and more detailed analyses of several related issues of hydrological 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.

Acknowledgements 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, Greece (EYDAP). 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. The constructive comments of an anonymous reviewer are gratefully appreciated. Thanks are due to two reviewers of an earlier version.

REFERENCES

Adenstedt, R. K. (1974) On large sample estimation for the mean of a stationary random sequence. *Ann. Statist.* **2**, 1095–1107.

Barnett, T. P., Hasselmann, K., Chelliah, M., Delworth, T., Hegerl, G., Jones, P., Rasmusson, E., Roeckner, E., Ropelewski, C., Santer, B. & Tett, S. (1999) Detection and attribution of recent climate change: a status report. *Bull. Am. Met. Soc.* **80**, 2631–2659.

Beran, J. (1994) *Statistics for Long-Memory Processes*, vol. 61 of Monographs on Statistics and Applied Probability. Chapman & Hall, New York, USA.

Bloomfield, P. (1992) Trends in global temperature. *Clim. Change* **21**, 1–16.

Bras, R. L. & Rodriguez-Iturbe, I. (1985) *Random Functions in Hydrology*. Addison-Wesley, Boston, Massachusetts, USA.

Carpenter, T. M. & Georgakakos, K. P. (2001) Assessment of Folsom lake response to historical and potential future climate scenarios: 1. Forecasting. *J. Hydrol.* **249**(1–4), 148–175.

Chow, V. T. (1964) Statistical and probabilistic analysis of hydrologic data. In: *Handbook of Hydrology* (ed. by V. T. Chow). McGraw-Hill, New York, USA.

Davey, M. K., Huddleston, M., Sperber, *et al.* (2002) STOIC: a study of coupled model climatology and variability in tropical ocean regions. *Clim. Dyn.* **18**, 403–420.

Eltahir, E. A. B. (1996) El Niño and the natural variability in the flow of the Nile River. *Water Resour. Res*. **32**(1) 131–137.

Haslett, J. & Raftery, A. E. (1989) Space–time modelling with long-memory dependence: assessing Ireland's wind power resource. *Appl. Statist.* **38**(1), 1–50.

Hirsch, R. M., Helsel, D. R., Cohn, T. A. & Gilroy, E. J. (1993) Statistical analysis of hydrologic data. In: *Handbook of Hydrology* (ed. by D. R. Maidment), Chapter 17, 17.1–17.55. McGraw-Hill, New York, USA.

Hosking, J. R. M. (1996) Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *J. Econometrics* **73**, 261–284.

Hurst, H. E. (1951) Long term storage capacities of reservoirs. *Trans. Am. Soc. Civil Engrs* **116**, 776–808.

Jones, P. D., Briffa, K. R., Barnett, T. P. & Tett, S. F. B (1998a) Millennial Temperature Reconstructions. IGBP PAGES/World Data Center-A for Paleoclimatology Data Contribution Series #1998-039, NOAA/NGDC Paleoclimatology Program, Boulder, Colorado, USA (ftp.ngdc.noaa.gov/paleo/contributions_by_author/jones1998/)

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

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

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

Koutsoyiannis, D. (2002) The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrol. Sci. J.* **47**(4), 573–596.

Koutsoyiannis, D. (2003) Internal report: http://www.itia.ntua.gr/getfile/537/2/2003HSJHurstSuppl.pdf

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

Lettenmaier, D. P., McCabe, G. & Stakhiv, E. Z. (1996) Global climate change: effect on hydrologic cycle. In: *Water Resources Handbook* (ed. by L. W. Mays), Chapter 29, 29.1–29.33. McGraw-Hill, New York, USA.

Mandelbrot, B. B. (1965) Une classe de processus stochastiques homothetiques a soi: application a la loi climatologique de H. E. Hurst. *C. R. Acad. Sci.* **260**, 3284–3277.

Mesa, O. J. & Poveda, G. (1993) The Hurst effect: the scale of fluctuation approach. *Water Resour. Res.* **29**(12), 3995–4002.

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

Nalbantis, I., Mamassis, N. & Koutsoyiannis, D. (1993) Le phénomène récent de sécheresse persistante et l'alimentation en eau de la cite d'Athènes. In: *Publications de l'Association Internationale de Climatologie* (6eme Colloque International de Climatologie, Thessaloniki, septembre 1993) (ed. par P. Maheras), **6**, 123–132, Association Internationale de Climatologie, Aix-en-Provence, France.

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

Przybylak, R. (2000) Temporal and spatial variation of surface air temperature over the period of instrumental observations in the Arctic. *Int. J. Climatol.* **20**, 587–614.

Radziejewski, M. & Kundzewicz, Z. W. (1997) Fractal analysis of flow of the river Warta. *J. Hydrol.* **200**, 280–294.

Salas, J. D. (1993) Analysis and modeling of hydrologic time series. In: *Handbook of Hydrology* (ed. by D. Maidment), Chapter 19, 19.1–19.72. McGraw-Hill, New York, USA.

Stedinger, J. R., Vogel, R. M. & Foufoula-Georgiou, E. (1993) Frequency analysis of extreme events. In: *Handbook of Hydrology* (ed. by D. R. Maidment), Chapter 18. McGraw-Hill, New York, USA.

Stephenson, D. B., Pavan, V. & Bojariu, R. (2000) Is the North Atlantic Oscillation a random walk? *Int. J. Climatol.* **20**, 1–18.

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

Vanmarcke, E. (1983) *Random Fields*. The MIT Press, Cambridge, Massachusetts, USA.

Von Storch, H., von Storch, J-S. & Müller, P. (2001) Noise in the climate system—ubiquitous, constitutive and concealing. In: *Mathematics Unlimited—2001 and Beyond* (ed. by B. Engquist & W. Schmid). Springer, Berlin, Germany.

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

Received 15 August 2002; accepted 17 September 2002