

**Uncertainty assessment of future hydroclimatic predictions:
A comparison of probabilistic and scenario-based approaches**

D. Koutsoyiannis

Department of Water Resources, National Technical University of Athens, Greece
Visiting scholar, Hydrologic Research Center, San Diego, CA., USA

A. Efstratiadis

Department of Water Resources, National Technical University of Athens, Greece

K. P. Georgakakos

Hydrologic Research Center, San Diego, CA., USA

Corresponding author:

K. P. Georgakakos

Hydrologic Research Center

12780 High Bluff Drive, Suite 250

San Diego, CA 92130-3017 USA

Tel: (858) 794-2726, Fax: (858) 792-2519

email: KGeorgakakos@hrc-lab.org

First draft: November 2005; Revision: June 2006

Submitted to Journal of Hydrometeorology

Abstract During the last decade, numerous studies have been carried out to predict future climate based on climatic models run on the global scale and fed by plausible scenarios about anthropogenic forcing to climate. Based on climatic model output, hydrologic models attempt then to predict future hydrologic regimes at regional scales. Much less systematic work has been done to estimate climatic uncertainty and to assess the climatic and hydrologic model outputs within an uncertainty perspective. In this study, a stochastic framework for future climatic uncertainty is proposed, based on the following lines: (1) climate is not constant but rather varying in time and expressed by the long-term (e.g. 30-year) time average of a natural process, defined on a fine scale; (2) the evolution of climate is represented as a stochastic process; (3) the distributional parameters of a process, marginal and dependence, are estimated from an available sample by statistical methods; (4) the climatic uncertainty is the result of at least two factors, the climatic variability and the uncertainty of parameter estimation; (5) a climatic process exhibits a scaling behavior, also known as long-range dependence or the Hurst phenomenon; (6) because of this dependence, the uncertainty limits of the future are affected by the available observations of the past. The last two lines differ from classical statistical considerations and produce uncertainty limits that eventually are much wider than those of classical statistics. A combination of analytical and Monte Carlo methods is developed to determine uncertainty limits for the nontrivial scaling case. The framework developed is applied with temperature, rainfall and runoff data from a catchment in Greece, for which data exist for about a century. The uncertainty limits are then superimposed onto deterministic projections up to 2050, obtained for several scenarios and climatic models combined with a hydrologic model. These projections indicate a significant increase of temperature in the future, beyond uncertainty bands, and no significant change of rainfall and runoff as they lie well within uncertainty limits.

1. Introduction

Projections of potential climate changes and ensuing land surface effects are typically based on the use of general circulation models (GCMs), with hypothesized scenarios (e.g. for the increase of CO₂ concentration, etc.). Such projections imply uncertainty, whose sources may be attributed to insufficient current understanding of climatic mechanisms, to inevitable weaknesses of numerical climatic and hydrologic models to represent processes and scales of interest, to complexity of processes and to unpredictability of causes (e.g. Strauss, 1993; Risbey and Stone, 1996; Mason et al., 1999; Anderson et al., 1999; Wang and Zwiers, 1999; Kharin and Zwiers, 2000; Shukla et al., 2000; Goddard et al., 2001). Some of these sources of uncertainty are intrinsic to the processes rather than related to our knowledge about the processes. As it is now well understood from studies of chaotic dynamical systems, even perfect knowledge of the system dynamics does not enable accurate predictions for the distant future and does not eliminate uncertainty (e.g. Tsonis, 1992).

Thus, the exact knowledge of future conditions, in other words the elimination of uncertainty, is probably impossible. Instead, a pragmatic target is the characterization (quantification) of uncertainty, which is very useful for decisions. Numerous studies attempt to approach this target by scenario-based analyses, which indicate possible trends of future climate evolution based on the assumed (scenarios) trends on forcings (e.g., IPCC, 2001; Giorgi et al., 2001; Cayan et al., 2001; Georgakakos and Smith, 2001; Dettinger et al., 2004). Few studies, however, have provided quantitative measures of uncertainty.

The mathematical tools appropriate to quantify uncertainty are provided by probability theory including statistics and stochastic processes. Such tools, widely used in the hydrologic practice, especially with the development of what has become known as the hydrologic statistics, are based on historical observed data. Their successful application emphasizes the fact that uncertainty exists even when the forcing conditions do not change, something that is sometimes missed under a scenario-based deterministic approach. The problem with probabilistic tools arises when forcing conditions (e.g. in the composition of the atmosphere

and in land uses) change significantly, so that the observations are no longer representative of the period of projection.

Another problem met in hydrologic statistics is the fact that it was built on the grounds of the classical statistical paradigm, which may not be appropriate to describe complex natural phenomena. Particularly, hydrologic statistical estimators of uncertainty are based on classical statistics, which assume independence of events. However, in the last decades evidence has accumulated that, not only are hydroclimatic phenomena dependent, but exhibit a peculiar temporal dependence structure known as scaling behavior, which is equivalent to the Hurst phenomenon or long-term persistence. Since the discovery of this behavior (Hurst, 1951), numerous studies of long time series affirm its omnipresence not only in hydrological or geophysical processes but also in economical, social and technological processes (for references see Kantelhardt et al., 2003; Koutsoyiannis, 2003). For a long time this behavior was regarded as a puzzle difficult to explain; for example Mesa and Poveda (1993) classified the Hurst phenomenon as one of the most important unsolved problems in hydrology and stated that “something quite dramatic must be happening from a physical point of view”. However, several explanations have been proposed ranging from conceptual (Klemeš, 1974; Montanari et al., 1999; Bhattachara et al., 1983; Koutsoyiannis, 2002) to more physically based ones (Klemeš, 1978; Beran, 1994, pp. 16-20). More recently, Koutsoyiannis (2006a) demonstrated using an extremely simplified climatic toy-model that this behavior can be explained by nonlinear climatic dynamics involving positive and negative feedbacks of the climatic system; also, Koutsoyiannis (2005b) showed that it may be the result of the maximum entropy principle applied to a spectrum of time scales. Furthermore, this scaling behavior harmonizes with the general recognition of a perpetually changing climate. Some have interpreted the Hurst behavior as a manifestation of nonstationarity; however, as explained in Koutsoyiannis (2006b) this is an inconsistent interpretation of long-term irregular climatic fluctuations. The idea of nonstationarity as a cause of the Hurst phenomenon is attributed to Klemeš (1974); however, a more attentive reading of his paper reveals that this is a misconception of his work; even though he used the term nonstationarity (in a loose context) he did point out that: “... the process employed in last set of experiments

can well be considered stationary". Indeed, his final models (processes with a Gaussian-distributed mean randomly varying from epoch to epoch) are stationary.

Adaptation of classical statistical estimators to the scaling paradigm is feasible (Koutsoyiannis, 2003). In this case, the obtained uncertainty band for a certain process becomes much wider than obtained by classical statistics, especially on multi-year time scales, on which climatic projections are done. The difference in results of typical statistical tests based on classical statistics on the one hand and those admitting long-term persistence on the other hand has been recently demonstrated by Cohn and Lins (2005) both theoretically as well as empirically based on the long term mean in northern hemisphere surface air temperature during the last century and a half. This difference emphasizes the importance of the choice of a (stochastic) model (or more generally of a modeling approach) to the quantification of uncertainty. The uncertainty related to the choice of a model (or approach) is hardly quantifiable per se. Other portions of uncertainty such as those due to natural variability, due to imperfect estimation of model parameters, and due to measurement errors can be quantified.

Even though the existence of scaling behavior has been known for more than half a century, its relationship to climatic uncertainty was only recently recognized as indicated from the above citations. Still there is a need to develop methodologies to streamline uncertainty estimation incorporating the scaling behavior. One may think of a stationary and a nonstationary setting of such methodologies. The stationary setting shall produce uncertainty measures assuming past and present forcing conditions and should include estimation of uncertainty for long as well as for short lead times. Estimation of uncertainty for short time scales is more difficult to achieve because the scaling behavior implies strong temporal dependence so that for short lead times future uncertainty limits are affected by the available observations of the past. The nonstationary setting would include in addition deterministic models to incorporate relationships between causes (forcings) and effects (hydroclimatology).

As a first step toward incorporating the scaling behavior in climatic uncertainty estimation, this paper is focused on a stationary setting (section 3). Although the emphasis in this methodology is on the scaling model, the classical statistical model is also used for

comparison. In addition and as an exploration step towards the nonstationary setting, comparisons of the uncertainty under stationary (e.g. present and past) conditions with the future climate trends suggested by GCM outputs are also provided. Such comparisons are very useful in order to assess how significant GCM projected hydroclimatic changes are, and how important it is to develop a nonstationary setting of the method. The methodology and the aforementioned ideas are illustrated based on a case study involving a catchment in Greece (described in section 2) in terms of three important hydroclimatic processes: temperature, rainfall and runoff. The uncertainty estimations and inter-comparison for the case study (section 4) are done for the first half of the 21st century for all three processes.

2. Test case and data sets

2.1 The Boeoticos Kephisos River basin

The case study is performed on an important basin in Greece, which is currently part of the water supply system of Athens and has a history, as regards hydraulic infrastructure and management, that extends backward at least 3500 years. This is the closed (i.e. without outlet to the sea) basin of the Boeoticos Kephisos River (Figure 1), with an area of 1955.6 km², mostly formed over a karstic subsurface. Owing to its importance for irrigation and water supply, data availability for the catchment extends for about 100 years (the longest data set in Greece) and modeling attempts with good performance have already been done on the hydrosystem (Rozos et al., 2004). The relatively long records have already made it possible to identify the scaling behavior of rainfall and runoff in this basin (Koutsoyiannis, 2003), and make the catchment ideal for a case study of uncertainty assessment.

The karstic subsurface of the watershed results in a significant groundwater yield directed to large springs in the basin or conducted to the sea. The hydrosystem serves multiple water uses. The surface resources are diverted to the neighboring Lake Hylike, one of the major water storage projects of Athens. Furthermore, important water supply and irrigation boreholes are located along the river and significantly affect the flow regime of the groundwater system.

2.2 Historical data sets

The long-term data set for the basin extends from 1908 to 2003 and comprises a flow record at the river outlet at the Karditsa station, rainfall observations in the raingage Aliartos and a temperature record at the same station; the station locations are shown in Figure 1. The original temperature record of Aliartos is for about 30 years (1967-97), but given the very high correlations of monthly temperature of Aliartos to that of Athens, it was possible to reconstruct a temperature record for the study period 1908-2003. The sample statistics of the three time series are shown in Table 1 on an annual basis. The convention of a hydrologic year (October of previous year to September of the current year) is used throughout this study; for easy comparison of runoff to rainfall, the flow record was converted into equivalent runoff depth with reference to the area of 1955.6 km² (this may differ from some earlier studies where an additional area of the Hylike lake basin is included).

In addition to this long-term but spatially sparse data set, a detailed hydrometeorological and hydrosystem control data set is available for a decade (1985-1994) at various sites of the basin. This, in addition to the record at Karditsa and Aliartos includes (1) point rainfall observations at 12 stations, which were integrated at five sub-basins using the Thiessen method (e.g. Koutsoyiannis and Xanthopoulos, 1999); (2) potential evapotranspiration, estimated by the Penman-Monteith method (e.g. Koutsoyiannis and Xanthopoulos, 1999); (3) water demand for water supply and irrigation, estimated according to the theoretical crop needs; (4) discharge records downstream of the main karstic springs; and (5) water table observations, with an average frequency of two per month.

2.3 Future projections

IPCC Scenarios

In addition to historical data, six GCM output data sets (past reconstructions and future projections) were used in this study for comparison. These correspond to three different climatic scenarios, summarized in Table 2, and come from three different GCMs, summarized in Table 3. A criterion we used to choose the GCM output data sets for the study, whose main

characteristics are shown in Table 4, was the extent of the output time series backward (to the past, in addition to future projections) so that we have a means to downscale the model outputs and compare the model performance based on the historical climatic evolution and especially its variability. These data sets of model runs are available on-line by the Data Distribution Centre of the Intergovernmental Panel of Climate Change (IPCC; http://ipcc-ddc.cru.uea.ac.uk/ddc_gcmdata.html).

Downscaling temperature and rainfall

A systematic technique for making inferences on small regional scales from coarser climate-model scales has been studied by Georgakakos (2003); however, a simpler technique was used here. For each scenario the model grid point nearest to the catchment was chosen for further analyses (see Figure 1). In addition, time series for the point second nearest to the catchment were also analyzed but these analyses are not included here as there is no essential difference from what was found based on the nearest point.

Time series of temperature and precipitation for the specified grid points were constructed and analyzed further; their time span extends to 2049 (some model outputs extend up to 2100 but the period beyond 2050 was not considered here). Two versions of each time series were constructed: the first, which is referred to as the initial series, is as extracted from the GCM outputs by use of the appropriate software packages and the second, which is referred to as the transformed series, is transformed to match the average and standard deviation for each month for the 30-year period 1960-1989; recall that in all model runs the period with historical input ends at 1989, so that from 1990 and beyond the output data constitute future projections. A linear transformation, different for each month, was used for temperature time series whereas for rainfall time series, to avoid negative values, a power transformation was used instead.

Construction of runoff time series

To estimate runoff under simulated forcings, and given the great extend of karst in the study basin, both surface and ground water processes must be modeled simultaneously; also, given that the basin is not in natural condition, the model must take into account both natural processes and anthropogenic influences on the catchment. To simulate the hydrological

processes of the Boeotikos Kephisos basin an advanced hydrologic modeling scheme, called Hydrogeios (Efstratiadis et al., 2005), was used, which can handle all these requirements. Hydrogeios is an improved version of an earlier modeling attempt (Rozos et al., 2004), which estimates the available water resources at characteristic sites of the river basin and the underlying aquifer. It is semi-distributed with parameters assigned on the basis of the physical characteristics of the watershed. The model requires geographical and hydrological input (precipitation and potential evapotranspiration, at a sub-basin scale), as well as management input (infrastructures, water needs, costs and priorities). The monthly time scale was chosen for this climatic study.

For setting up the hydrologic model, the detailed data set described in section 2.2 was split into two subsets (or subperiods), consisting of the calibration subperiod of six years (October 1984-September 1990) and the validation subperiod of four years (October 1990-September 1994). In a second step the model was run for the period 1908-2003 with the long term historical data set in order to assess its behavior in comparison with the historical runoff time series. To acquire detailed input for the model run for this period, the available long term time series were downscaled as described below. In a third step, the same downscaling procedure was applied with GCM output time series and the hydrologic model was run to obtain future runoff time series according to each of the climatic scenarios.

The MOVE.1 technique (Hirsh et al., 1993) was used for the downscaling of the rainfall time series, i.e. the derivation of rainfall time series for each of the five sub-basins on the basis of the point rainfall time series at Aliartos. The technique preserves both mean and variance, and is thus superior to linear regression models for this study, given the importance of variance in uncertainty estimation. The same method was used for downscaling temperature.

To derive potential evaporation time series from temperature, especially for the GCM scenarios, it was first attempted to use incoming solar radiation produced by the models. Inspection of the climatic time series of solar radiation on a 30-year time scale, as obtained from the GCM outputs for the models and scenarios described above, showed that the climatic variability of radiation is rather negligible and that different models, rather than

different scenarios, produce different radiation (probably as a result of different definition or representation of the quantity in each model). Given these disagreements among models, the absence of measured solar radiation data, which would enable standardization of synthetic series, and the negligible variation of climatic solar radiation indicated by the models, eventually it was preferred not to use the GCM output for solar radiation. Rather, to estimate the potential evaporation in a downscaling framework, the following approximate relationship (Koutsoyiannis and Xanthopoulos, 1999, p. 223) was used,

$$E_p = \frac{a S_0 - b}{1 - c T} \quad (1)$$

where S_0 is the incoming solar radiation at the top of the atmosphere (not affected by changes in the atmosphere composition), T the surface temperature and a , b and c parameters that are estimated by minimizing the squared error from Penman-Montieth potential evaporation (for the period of the detailed data set). As shown in Koutsoyiannis and Xanthopoulos (1999), if there exist data to calibrate the parameters of this equation, its accuracy is very high (with a determination coefficient value of 0.99).

For the water demand, which is also a necessary input to the hydrologic model, assumptions were made considering the history of agricultural and urban development in the area. Specifically, it is assumed that in 1984 the demand reached a saturation value which is maintained in the 21st century. During 1900-1984, a 2% increase of the total demand per year was assumed. Before 1900 (for scenarios MP01GG01 and MP01GS01), zero demand was assumed (prior to 1900 the Copais plain was a lake).

Despite the unavoidable simplifications regarding hydrological input, especially the downscaling of the areal precipitation, the model performance for the entire 1908-2003 period based on historical input was very satisfactory, comparable to that of the calibration period (1984-1990). The coefficient of efficiency (defined in (25)) for the runoff at the outlet was 75% and 68%, on a monthly and annual basis, respectively. These scores reflect a generally good model performance with some isolated significant departures of the simulated from the observed values for a limited number of flood events, for which the lack of spatial rainfall

reduced performance. The good model performance with historical input renders the model capable for use with GCM projections of runoff given the projections of temperature and rainfall. Thus, for each of the GCM scenarios, it was possible to produce a runoff scenario by running the hydrologic model.

3. Probabilistic quantification of hydroclimatic uncertainty

3.1 Description of uncertainty

The quantification of uncertainty is done using notions of probability and statistics, which can be found in mathematical statistics books such as Papoulis (1991) and Bickel and Doksum (2000); a summary of notions compiled for this study, along with clarifications for a few required modifications can be found in Koutsoyiannis et al. (2006). The most important modification is that classical statistics are based on the assumption that a sample is a sequence of independent identically distributed random variables, whereas in our framework the variables representing a climatic process are assumed dependent in time.

Let X_i represent a hydrometeorological quantity at the annual scale at time i . We use the convention that $i = 0$ corresponds to present, $i < 0$ to the past and $i > 0$ to the future. Furthermore, we assume that there is a record of n observations of the present and past, which we write as a vector $\mathbf{x}_{0,n} = [x_0, \dots, x_{1-n}]'$ (where the prime is used to denote the transpose of a vector or matrix). We recall from statistical theory that each observation x_i represents a realization of a random variable X_i , so that $\mathbf{x}_{0,n}$ is a realization of a vector of identically distributed random variables $\mathbf{X}_{0,n} = [X_0, \dots, X_{1-n}]'$ (notice the upper and lower case symbols used for random variables and values thereof, respectively, and the arrangement of the sample members and observations from the latest to the earliest). In natural macroscale phenomena this realization is unique, whereas the theory of stochastic processes implies infinite realizations of the process of infinite length. This is not a contradiction: a stochastic process is a mathematical model whereas our observations are for a real world process with unique evolution. We use the model to predict future behavior and we infer the model structure and parameters from our observations $\mathbf{x}_{0,n}$ by imposing some fundamental postulates such as those

of stationarity and ergodicity (for justification and interpretation of the stationarity postulate see Koutsoyiannis, 2006b).

It is sometimes meant that climate is represented by a parameter β , such as the mean annual rainfall at a certain location. The numerical value of this parameter is unknown but it can be estimated (approximated) by the observed sample mean $\bar{x} = (x_0 + \dots + x_{1-n}) / n$. The smaller the sample size n , the higher the uncertainty about the true value β , which can be called parameter uncertainty. This uncertainty is represented by a pair of statistics (random variables) U (upper) and L (lower) called the interval estimators of β for a chosen confidence coefficient α . These variables have the property $P\{L < \beta < U\} = \alpha$. Their observed sample estimates l, u define the interval estimate (l, u) of β , which represents the quantified measure of parameter uncertainty.

In a few special cases analytical derivation of confidence limits is possible. In all other cases the determination of confidence limits requires numerical methods such as Monte Carlo simulation. By construction, Monte Carlo simulation assumes a fully known model (along with its parameters) and therefore can readily produce prediction limits of random variables rather than confidence limits of parameters (for clarification of the difference of prediction limits and confidence limits see Papoulis, 1991, and Koutsoyiannis et al. 2006). The calculation of confidence limits of parameters is much more difficult than that of prediction limits and requires repetitive simulations with different parameters (in a sort of inverse problem setting). Here an approximate Monte Carlo simulation technique was used, which is general enough to yield confidence limits of any distributional parameter (marginal or dependence) or any combination of parameters for any marginal distribution and dependence structure of the stochastic process of interest, provided that there exists a stochastic generator of this process (see Koutsoyiannis and Kozanis, 2005, for details).

The notion of a climatic parameter (a number constant in time) is, however, not very helpful in climatic studies. When we speak of climatic variability in time we indirectly assume that climate is better represented as a time varying quantity (i.e. a stochastic process, a time indexed family of random variables) rather than as a constant parameter. Specifically, it has been common practice in climatology as well as in hydrology to consider as climate the

time average of a certain quantity over a certain number of years. In this case, the climate at any time (year) i can be represented as the discrete time stochastic process $X_i^{(k)}$ (also termed a stochastic sequence or a digital process) defined in terms of the annual process X_i as

$$X_i^{(k)} := (X_i + \dots + X_{i-k+1})/k \quad (2)$$

This notation implies that $X_i^{(1)} \equiv X_i$ (that is, a superscript equal to one can be omitted) furthermore, $X_i^{(k)}$ is an ‘overlapping’ moving average if $i = \dots, -1, 0, 1, 2, \dots$ and a ‘non-overlapping’ average if $i = \dots, -k, 0, k, 2k, \dots$ (we use both options in the following sections). Typically, the convention $k = 30$ is used to standardize the climatic time scale (number of years); a less frequently used value is $k = 10$ (for example, IPCC provides records of observed global climate on a 30-year, as well as a 10-year basis; <http://ipcc-ddc.cru.uea.ac.uk/obs/index.html>). Here, unless otherwise stated, it is assumed that $k = 30$.

The quantification of uncertainty with the notion of a climatic process is more complex than in the case of a climatic parameter. Here, in addition to parameter uncertainty, we have also the uncertainty due to (natural) temporal variability. As a stochastic process is a time indexed sequence of random variables, there is a range of characterizations that may be employed depending on the order of the joint distribution we are willing to consider for the characterization (e.g., first order characterization by the one-dimensional distribution at a single time for all times, second order characterization for the two dimensional joint distribution between any two times for all times, etc.). For a first order characterization at any scale k , let $F^{(k)}(y)$ be the distribution function of each of $X_i^{(k)} \equiv Y$. To express quantitatively the uncertainty due to temporal variability we choose a confidence coefficient α' and we find the distribution quantiles y_b (below) and y_a (above) such that

$$P\{y_b < X_i^{(k)} < y_a\} = \alpha' \quad (3)$$

(see also Figure 2). If $F^{(k)}(x)$ is precisely known and its inverse is $(F^{(k)})^{-1}$, then y_b and y_a can be determined from (e.g. Papoulis, 1990)

$$y_b = (F^{(k)})^{-1}(b), \quad y_a = (F^{(k)})^{-1}(a), \quad b = (1 - \alpha')/2, \quad a = (1 + \alpha')/2 \quad (4)$$

Precise knowledge of $F^{(k)}$ means that both its mathematical expression (model) and its parameters are known. In practice, none of these are known. The model is simply hypothesized, based on exploration of available observations, on previous experience and, in the best case, on theoretical reasoning. The parameters are estimated from the available historical record. Thus, on top of the uncertainty due to temporal variability, we have additional uncertainty due to parameter estimation. Because both y_b and y_a are not variables but parameters (dependent on the distributional parameters), their uncertainty is expressed as discussed above about parameter uncertainty. In this case we need to choose another confidence coefficient α (not necessarily equal to α' but in this study, unless stated differently, it is assumed $\alpha = \alpha' = 0.95$). Thus, the uncertainty of y_b is expressed by the confidence interval estimate $(l(y_b), u(y_b))$ and that of y_a by the confidence interval estimate $(l(y_a), u(y_a))$. Consequently, the total uncertainty can be expressed by the interval $(l(y_b), u(y_a))$ (see explanation in Figure 2).

Although uncertainty is also introduced by measurement errors, this is neglected here assuming that it is much smaller compared to the other two sources of uncertainty.

3.2 The scaling property

In classical statistics, if $\sigma^{(k)} := \text{StD}[X_i^{(k)}]$ denotes the standard deviation of the random variable $X_i^{(k)}$ at scale k and any time i , we have the fundamental law

$$\sigma^{(k)} = \frac{\sigma}{\sqrt{k}} \quad (5)$$

where $\sigma \equiv \sigma^{(1)}$ is the standard deviation of each of X_i . Given a time series of sufficient length n , we can test in a simple way whether this law is fulfilled or not. To this aim, from the observed time series x_i we form the time series $x_i^{(k)}$ with $k = 1, 2, \dots$, (up to about $n/10$, so that the resulting time series have at least ten elements) and $i = 0, -k, -2k, -3k, \dots$, and from the time series of each k we calculate the estimate $s^{(k)}$ of the standard deviation $\sigma^{(k)}$. Plotting $s^{(k)}$ versus k (preferably on a logarithmic plot) we can test graphically the validity of the statistical law (5). Such plots for the test time series (temperature, rainfall and runoff in Boeoticos

Kephisos) are given in Figure 3. The non fulfillment of law (5) is clear and all three plots of the real world time series suggest a more general law of the form

$$\sigma^{(k)} = \frac{\sigma}{k^{1-H}} \quad (6)$$

where H is a constant between 0.5 and 1. As shown in Figure 3, the value of H ranges between 0.64 and 0.79 in the three series.

A stochastic process that has the property (6) is mathematically feasible and can be called a simple scaling stochastic process (SSS process; else it is known as the stationary intervals of a self-similar process, or a fractional Gaussian noise if its distribution function is normal). In fact, (6) can serve as a definition of an SSS process, sufficient for the purpose of this paper (see also Koutsoyiannis, 2002, 2003; for a definition in continuous time see Beran, 1994). The statistical behavior expressed by equation (6) is in fact the scaling behavior and the constant H is the Hurst exponent. It is stressed that the SSS process is stationary.

Another way to demonstrate the departure of the natural behavior from the classical statistical model is through very low and very high flows over different scales. Such a demonstration is depicted in Figure 4 for the Boeoticos Kephisos 96-year runoff record. The record was aggregated from scales 1 to 10 (years) and for each scale the minimum and maximum value was found. For these values, the return periods were calculated assuming that the distribution of $X_t^{(k)}$ is normal and that the standard deviation over scale k is given by (5) and (6) for the classical and SSS model, respectively. As shown in Figure 4, the classical model renders return periods reaching 100 000 years for large scales. The reason for this is that some dry (or wet) years are in reality clustered in multiyear periods but this is quite unlikely according to the classical model. In nature, however, a clustering behavior is very common and this is well represented by the SSS model, which for all scales renders return periods around 100 years to the minima and maxima of all scales, as expected for a record with length of about 100 years.

3.3 Unconditional uncertainty

Let y_b and y_a be the b and a distribution quantiles of $X_i^{(k)}$ as in (4); these describe the uncertainty due to temporal variability to a first order characterization. We wish to find the total uncertainty by incorporating the parameter uncertainty as described in section 3.1. We assume for convenience that the process of interest X_i is normally distributed; this can be justified for the annual scale and was empirically verified (Koutsoyiannis, 2005a). A fortiori, $X_i^{(k)}$ is also normally distributed. Strictly speaking, in all cases the distribution should be assumed as truncated normal because all variables are bounded from below (rainfall and runoff by zero, temperature by absolute zero; Koutsoyiannis, 2005a). The analytical handling of the truncated normal distribution is difficult, so to simplify calculations the truncated normal distribution is approximated by a normal distribution and if needed a truncation is done a posteriori to confidence limits estimated numerically.

If μ and σ are the parameters of the (marginal) normal distribution of the process at scale 1, then in classical statistics, by virtue of (5) the quantities whose confidence limits are to be estimated are

$$y_b^{(k)} = \mu + \zeta_b \sigma / \sqrt{k}, \quad (7)$$

where ζ_b is the b quantile of the standard normal distribution, and $y_a^{(k)}$, whose expression can be derived from (7) by replacing the subscript b with a . These quantities combine the uncertainty in the estimation of both μ and σ (which are estimated by the observed sample mean $\bar{x} \equiv x_0^{(n)}$ and the standard deviation s). Similarly, by virtue of (6), the corresponding quantities for the SSS case are

$$y_b^{(k)} = \mu + \zeta_b \sigma / k^{1-H} \quad (8)$$

and $y_a^{(k)}$, and combine the uncertainty in the estimation of μ , σ and H .

Combining known results in hydrological statistics (e.g., Stedinger et al., 1993, p. 18.30) and equation (5), we can easily obtain approximate confidence limits of $y_b^{(k)}$, which for the classical case are

$$u(y_b^{(k)}), l(y_b^{(k)}) = x_0^{(n)} + \zeta_b s / \sqrt{k} \pm \zeta_{(1+\alpha)/2} \varepsilon_b s \quad (9)$$

where

$$\varepsilon_b = \frac{1}{\sqrt{n}} \sqrt{1 + \frac{\zeta_b^2}{2k}} \quad (10)$$

Generalization of these equations for the SSS case with known H has been proposed in Koutsoyiannis (2003). With a small adaptation, these confidence limits are expressed as

$$u(y_b^{(k)}), l(y_b^{(k)}) = x_0^{(n)} + \zeta_b s / k^{1-H} \pm \zeta_{(1+\alpha)/2} \varepsilon_b s \quad (11)$$

where

$$\varepsilon_b = \frac{1}{n^{1-H}} \sqrt{1 + \frac{\varphi(n, H)}{2 n^{2H-1}} \left(\frac{\zeta_b}{k^{1-H}} \right)^2}, \quad \varphi(n, H) = (0.1 n + 0.8)^{0.088(4H^2 - 1)^2} \quad (12)$$

It can be easily verified that equations (9) and (10) are special cases of (11) and (12), respectively, obtained when $H = 0.5$.

The analytical derivation of confidence limits for the realistic case where H is unknown and is estimated from the sample may be intractable. In this case, the Monte Carlo method can be used based on the quantities defined in equations (7) and (8) (the analytical results (9)-(12) are not applicable in this case). The Monte Carlo confidence limits are illustrated in Figure 5, which also compares classical confidence limits to those of the SSS case; the sample statistics used to calculate the confidence limits are those of runoff shown in Table 1. Although in this figure only the Monte Carlo confidence limits are plotted, where applicable the theoretical ones were also evaluated and were found to be virtually indistinguishable from the Monte Carlo ones. Specifically, for the classical case, the theoretical confidence limits evaluated from (9) are indistinguishable from those marked MCCL/classical in Figure 5, and the theoretical confidence limits evaluated from (11) are indistinguishable from those marked MCCL/SSS 1 (for known H).

The upper panel of Figure 5 refers to the annual quantities (i.e. $k = 1$). In this case both the classical and the SSS models result in the same amount of uncertainty due to variability (the

point estimates marked as PE in Figure 5 are the same). However, there are significant differences in the parameter uncertainty. The confidence band in the classical case is narrow and becomes much wider in the SSS case with known H (marked MCCL/SSS 1 in Figure 5) and even wider in the case of unknown H (marked MCCL/SSS 2).

More interesting is the lower panel of Figure 5, which refers to the climatic quantities (i.e. $k = 30$). The low variability and uncertainty in the classical model is depicted as a narrow, almost horizontal, band in Figure 5. Here, the SSS model, in addition to the higher parameter uncertainty, results in uncertainty due to variability much wider than in the classical one (obviously this is implied by the term k^{1-H} in (11) as compared to the term \sqrt{k} in (9)). As a result, while in the classical model the total uncertainty (width of the interval) is about 50% of the mean, in the SSS case it becomes about 200% of the mean, or four times larger. In addition, according to the SSS model, the uncertainty at the 30-year scale is only slightly lower than that of the annual scale (the latter is 268% of the mean). This contrasts to the classical model, which yields significant reduction as we proceed from the annual to the climatic scale.

A comparison of the uncertainty of climatic values of rainfall and runoff is depicted in Figure 6. The lower panel of this, which refers to runoff, is identical to the lower panel of Figure 5 except in the vertical scale, which was changed to agree with that of the upper panel of Figure 6; the upper panel refers to rainfall and was constructed in the same manner as that of runoff. The comparison of the two panels in terms of the results of the classical model suggests that the uncertainty in runoff is lower than that in rainfall. This is an inconsistency of the classical model as it is well known that the variability of rainfall is magnified under the rainfall-runoff transformation, provided that the same spatial scale is used for both processes. In the SSS model the inconsistency is remedied as the uncertainty of runoff is higher than that of rainfall.

3.4 Conditional uncertainty

The confidence limits obtained in section 3.3 describe the uncertainty on a long lead time (theoretically for lead time tending to infinity, so that the influence of the past information

becomes negligible) and need to be adapted for the short lead time by taking into account the information from observations. The adaptation could be done either by modifying the Monte Carlo method already used or by deriving analytical equations, which are then evaluated by the Monte Carlo simulation technique mentioned earlier (Koutsoyiannis and Kozanis, 2005). Here we preferred the second option because analytical equations provide some insight into the factors influencing uncertainty.

The adaptation is very simple in the case of the classical model. For a lead time longer than the climatic time scale ($i \geq k$) the unconditional confidence limits apply. For a shorter lead time ($i < k$) the climate, conditional on the present and past information, could be written

$$X_i^{(k)} | \mathbf{x}_{0,n} := (X_i + \dots + X_1 + x_0 + \dots + x_{i-k+1})/k = \frac{i}{k} X_i^{(i)} + \left(1 - \frac{i}{k}\right) x_0^{(k-i)} \quad (13)$$

where it was assumed that the available observation period is longer than the climatic time scale ($n > k$). Because in the classical model $X_i^{(i)}$ is independent from $X_0^{(k-i)}$, it suffices to calculate confidence limits for the quantities

$$y_{b,i}^{(k)} = \frac{i}{k} y_b^{(k)} + \left(1 - \frac{i}{k}\right) x_0^{(k-i)} = \frac{i}{k} \left(\mu + \zeta_b \frac{\sigma}{\sqrt{k}}\right) + \left(1 - \frac{i}{k}\right) x_0^{(k-i)} \quad (14)$$

and $y_{a,i}^{(k)}$. We note that the last term of (14) is a constant known without uncertainty (as we have assumed that the observation errors are negligible). Thus, the conditional confidence limits are simple linear transformations of the unconditional ones.

In the SSS case, the situation is much more complex because of the temporal dependence. As shown in section 3.3, the dependence results in a dramatic increase of the total (but unconditional) uncertainty, especially on the climatic time scale. At the same time the dependence, combined with observation of the past, apparently has a reducing effect to the total uncertainty on a conditional setting. For long lead time ($i \geq k$), the climatic quantity of interest is

$$X_i^{(k)} | \mathbf{x}_{0,n} = [(X_i + \dots + X_{i-k+1})/k] | \mathbf{x}_{0,n} = \frac{1}{k} \sum_{j=i-k+1}^i X_j | \mathbf{x}_{0,n} \quad (15)$$

and can be also be written as

$$X_i^{(k)}|\mathbf{x}_{0,n} = \frac{1}{k} \sum_{j=1}^i X_j|\mathbf{x}_{0,n} - \frac{1}{k} \sum_{j=1}^{i-k} X_j|\mathbf{x}_{0,n} = \frac{i}{k} X_i^{(i)}|\mathbf{x}_{0,n} + \left(1 - \frac{i}{k}\right) X_{i-k}^{(i-k)}|\mathbf{x}_{0,n} \quad (16)$$

For shorter lead time ($i < k$), the quantity of interest is (cf. equation (13))

$$X_i^{(k)}|\mathbf{x}_{0,n} = \frac{1}{k} \sum_{j=1}^i X_j|\mathbf{x}_{0,n} + \left(1 - \frac{i}{k}\right) x_0^{(k-i)} = \frac{i}{k} X_i^{(i)}|\mathbf{x}_{0,n} + \left(1 - \frac{i}{k}\right) x_0^{(k-i)} \quad (17)$$

Thus, to assess the confidence limits of $X_i^{(k)}|\mathbf{x}_{0,n}$ we need to calculate the expected value and standard deviation of $X_i^{(k)}$ conditional on the present and past being observed.

Based on the above decompositions, it is shown in the Appendix that the conditional climatic expectation is given by the following approximation (exact equations are also given in the Appendix but the approximation is good for our purpose):

$$E[X_i^{(k)}|\mathbf{x}_{0,n}] = \left[\frac{i}{k} \varphi_{i,n}(H) + \left(1 - \frac{i}{k}\right) \varphi_{i-k,n}(H) \right] \mu + \left\{ \frac{i}{k} [1 - \varphi_{i,n}(H)] + \left(1 - \frac{i}{k}\right) [1 - \varphi_{i-k,n}(H)] \right\} x_0^{(n)}, \quad i \geq k \quad (18)$$

$$E[X_i^{(k)}|\mathbf{x}_{0,n}] = \frac{i}{k} \varphi_{i,n}(H) \mu + \frac{i}{k} [1 - \varphi_{i,n}(H)] x_0^{(n)} + \left(1 - \frac{i}{k}\right) x_0^{(k-i)}, \quad i \leq k \quad (19)$$

where $\varphi_{i,n}(H) := 1 - \xi_{i,n}' \boldsymbol{\tau}_n^{-1} \mathbf{e}_n$; \mathbf{e}_n is a size n vector with all its elements equal to 1; $\boldsymbol{\tau}_n := \text{Cov}[\mathbf{X}_{i,n}, \mathbf{X}_{i,n}]/\sigma^2$ is the symmetric size n correlation matrix of the process whose $(i, i \pm j)$ item is the lag j autocorrelation of the process of interest at the scale 1 (annual), i.e. $\rho_j := \text{Cov}[X_i, X_{i+j}]/\sigma^2$; and $\xi_{i,n}$ is a size n vector defined as

$$\xi_{i,n} := \frac{1}{i} \sum_{j=1}^i \boldsymbol{\rho}_{j,n} \quad (20)$$

whereas $\boldsymbol{\rho}_{j,n} := [\rho_j, \dots, \rho_{j+n-1}]'$. Recall that in the SSS model the lag j autocorrelation is (e.g. Koutsoyiannis, 2002)

$$\rho_j^{(k)} = \rho_j = (1/2) [(|j+1|)^{2H} + (|j-1|)^{2H}] - |j|^{2H} \quad (21)$$

Clearly then, $\varphi_{i,n}(H)$ is a function of the Hurst coefficient H , a lag i and the observation length n . A plot of $\varphi_{i,n}(H)$ for varying i and H and a fixed $n = 100$, which is close to the sample sizes used in this study, is shown in Figure 7 (left panel). Approximate analytical relationships for $\varphi_{i,100}(H)$ are given in the Appendix. As per the implementation of (19) in a Monte Carlo framework to obtain confidence limits, it should be noted that only i , k , $x_0^{(n)}$ and $x_0^{(k-i)}$ are known; all other parameters that appear are unknown.

Figure 8 depicts the evolution of conditional mean of the Boeoticos Kephisos runoff both at the annual and the climatic scale and also provides a comparison of the approximation (19) with the exact equation given in the Appendix (equation (A.4)). It can be observed that the differences are small and that in a 30-year period the conditional mean approaches the true mean which to construct Figure 8 was assumed to be equal to the sample mean. The later assumption is only for demonstration purposes and used only for Figure 8. (The Monte Carlo simulation technique used in this study is consistent with the fact that the true mean is unknown; in fact, in an SSS process the sample mean can depart significantly from the true mean). The influences of the three terms involved in the calculation of the conditional mean are depicted in Figure 9. The first term, which is the coefficient of the true mean μ in (19) is an increasing function of the lead time i . The second term, which is the coefficient of the sample mean $x_0^{(n)}$ in (19), is an increasing function of the lead time i up to $i = k = 30$ and then decreases. The third term, which is the constant term (the last term in (19)) divided by the true mean, is obviously a decreasing function of i and vanishes off at $i = k = 30$. For $i > k$, the sum of the first two terms is 1 and at about $i = 2k = 60$ both terms are equal to 0.50. Even for lead times as high as 100, the influence of the sample average, which is known, is significant and the influence of the true mean, which is unknown, is smaller than 60%. Certainly, this has an attenuating effect to the width of the confidence band.

Furthermore, as derived in the Appendix, the conditional standard deviation of $X_i^{(k)}$ can be approximated by

$$\text{StD}[X_i^{(k)} | \mathbf{x}_{0,n}] = k^{H-1} \sigma \sqrt{\psi_{i/k}(H)}, \quad i \geq k \quad (22)$$

$$\text{StD}[X_i^{(k)}|\mathbf{x}_{0,n}] = \frac{i^H}{k} \sigma \sqrt{\psi_1(H)}, \quad i \leq k \quad (23)$$

where $\psi_i(H) := \lim_{n \rightarrow \infty} \psi_{i,n}(H)$ and $\psi_{i,n}(H) := 1 - \boldsymbol{\rho}_{i,n}' \boldsymbol{\tau}_n^{-1} \boldsymbol{\rho}_{i,n}$. Clearly, $\psi_i(H)$ is a function of the Hurst coefficient H and a lag i , as plotted in Figure 7 (right panel); an approximate analytical expression is given in the Appendix. Although (22) is precisely valid for integral values if l/k , its use with fractional values is possible and can give approximate interpolations.

The variation with the lead time of the conditional standard deviation (divided by the unconditional standard deviation) of the Boeotikos Kephisos runoff at the climatic scale time is depicted in Figure 9 (along with that of the terms involved in the calculation of the conditional mean, discussed earlier). It can be observed that the conditional standard deviation increases rapidly with the lead time, reaching 85% at $i = k = 30$. Then it increases more gradually and at a lead time 100 it becomes about 93% of the unconditional value.

The parameters whose conditional confidence limits are sought are the distribution quantiles

$$y_{b,i}^{(k)} = E[X_i^{(k)}|\mathbf{x}_{0,n}] + \zeta_b \text{StD}[X_i^{(k)}|\mathbf{x}_{0,n}] \quad (24)$$

and $y_{a,i}^{(k)}$, where the quantities $E[X_i^{(k)}|\mathbf{x}_{0,n}]$ and $\text{StD}[X_i^{(k)}|\mathbf{x}_{0,n}]$ are evaluated from (18)-(23); the first is a linear function of μ and a nonlinear function of H , the second is a linear function of σ and a nonlinear function of H , and both depend also on the lead time i , the climatic time scale k , the observed mean $x_0^{(n)}$ of the entire observation period and the observed mean $x_0^{(k-i)}$ of the most recent observations. Assuming that H is unknown, the nonlinear functions of it involved in the estimation of conditional expectation and standard deviation increase variability on a Monte Carlo calculation framework. For this reason, for a small H and long lead time, it may be anticipated that the conditional confidence limits may come up wider than the unconditional ones; in such a case the unconditional confidence limits should be used instead.

The results of the application of the method for all three processes examined here (temperature, rainfall and runoff with sample statistics as in Table 1) are depicted in Figure 10 for a horizon up to 2049. For comparison the classical conditional confidence limits have been

plotted, along with the SSS confidence limits. It is observed that the SSS confidence bands are much wider (about 2.3 to 3 times) than the classical ones.

4. Comparisons and validations

4.1 Performance of GCM output time series in the past

Even though the main target of this section is to inspect and compare the developed stochastic uncertainty framework of future climate with the deterministic GCM future projections, it is useful to compare first the performance of GCM past climate traces with observed climate. The performance of the GCM output time series in the past is indicated in Table 5 in terms of several performance indicators as well as in Figures 11-12 in graphical depictions. A first observation in Table 5 is the notable positive bias in all initial series of temperature and the significant negative bias on the rainfall. These, however, were expected in the three out of four cases because the grid points lie to the south of the catchment, in an area with higher temperature and lower rainfall; in the CCCma_A2 and B2 case this was rather unexpected as the grid point lies in the northern border of the catchment. However, the most significant drawback of all model outputs is the underdispersion (negative overdispersion) both in temperature and rainfall. This is obvious both in Table 5 as well as in Figure 11 and has three components. First, in each month the standard deviation of temperature and rainfall (Figure 11, right column) is much lower than the historical one (in some case less than half). Second, the periodic variation through the months within the year is lower than the historical; this is better visible through the profiles of the means (Figure 11 left column). And third, the annual and interannual variability is again lower than the historical as understood from the high negative overdispersion values in Table 5 at the annual scale, the small Hurst coefficients in Table 5 and the flat climatic plots in Figure 12. There is one exception, the temperature time series of MP01GG01, which as shown in Figure 12 is not flat and yields a Hurst coefficient 0.76, even higher than the historical 0.72. The plots in Figure 12, however, simultaneously show that the evolution in the MP01GG01 temperature series is opposite to the historical one (overyear trends just opposite to the historical ones). In Table 5 it is also shown that the

temperature time series of the HADCM3_A2 and B2 cases has a large Hurst coefficient (0.86). This however is an artifact resulting from small sample (40 years in HADCM3_A2 and B2), so it does not represent a high over-annual variability (see Figure 12).

The bias and the underdispersion caused by the former two reasons are corrected or significantly improved in the transformed series both at the monthly scale and at the annual scale. However, it is impossible to rectify this on the over-annual climatic scale.

Other measures of performance are the coefficients of determination (C_d) and efficiency (C_e) between model time series (x_i') and historical ones (x_i). These are defined respectively by

$$C_d = \frac{\left[\sum_{i=1}^n (x_i - \bar{x}) (x_i' - \bar{x}') \right]^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (x_i' - \bar{x}')^2}, \quad C_e = 1 - \frac{\sum_{i=1}^n (x_i' - x_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (25)$$

where n is the length of the time series and \bar{x} and \bar{x}' are their averages. The former is a relative fitting measure between the time series x_i and x_i' postulating a linear regression between them ($x_i' = a + b x_i$) and its values are between 0 and 1. The latter is a fitting measure over an equality postulate and its values are smaller than 1 and can be also negative in case of poor agreement. Obviously, the latter is more appropriate in our case as the two time series ideally should be identical, not just linearly related.

The values of these indicators are shown in Table 5. For the monthly temperature time series the two coefficients are good enough in the initial series and become almost perfect in the transformed ones. This, however, is not seen in rainfall, where, because of the high variation, the fitting measures are poor even after transformation. At the annual scale, these coefficients are very poor for both temperature and rainfall. Finally, at the climatic (30-year) scale, Figure 12 shows that the GCM outputs did not capture the general climatic trends in the past. In the intermediate period 1990-2003, which for the GCM model runs was 'future' but now is past (observed), the model projections do not agree with observations (except in ECHAM4/OPYC3).

The performance of the combination of GCM and hydrologic model outputs at the annual scale is given in Table 6. Initially, it can be observed that the hydrological model alone with historical temperature and rainfall performs well, as indicated from the low bias and overdispersion and the high coefficients of determination and efficiency in this case (first row in Table 6). However, the hydrological model with temperature and rainfall data from GCM outputs, transformed as described above, does not reproduce the past (high negative bias, high underdispersion, low coefficients of determination and efficiency). It is interesting, however, that the Hurst coefficients of runoff from all GCM scenarios are high, even though those of the rainfall input are low. This can be explained as the combined effect of several factors such as (a) the dominance of groundwater processes in runoff, which introduces strong temporal dependence, (b) the increasing groundwater abstractions through the historical period, which result in a monotonic downward trend, (c) the enhancement of this trend by the temperature increase predicted by GCM for the last years, and (d) the significant underdispersion of the resulting runoff at the annual scale, combined with smaller underdispersion at larger scales because of the effect of the downward trends.

4.2 GCM future projections

For the future, the GCM rainfall falls within the SSS uncertainty limits for the whole examined period up to 2049 (Figure 12; some departures in the period 1990-2000 just indicate that the projections of these models were not validated by the historical evolution). This means that traces such as the ones projected by the GCM can be readily obtained by stochastic simulation assuming stationarity. This is not the case, however, for temperature, where all model projections agree that sooner or later the temperature will depart from the uncertainty limits. The future runoff estimated by the hydrological model for GCM future temperature and rainfall appears to be almost constant for all scenarios and models. All climatic traces of runoff are concentrated into a narrow band around the current climatic value. This more or less harmonizes with some earlier studies; for example Georgakakos and Smith (2001) in a comprehensive climatic study for the United States based on GCM output concluded that the soil moisture deficits (which are closely related to runoff) in most areas are

of the same order of magnitude as the soil moisture field of standard deviations arising from historical natural variability. The stochastic method developed in this study suggests that the variability and uncertainty of runoff, even assuming natural conditions, is in fact very high as becomes evident from the confidence band in Figure 12, whereas GCM projections fail to reproduce even the natural variability.

4.3 Further validation

One may argue that the 20th century data that were used in this study are already affected by anthropogenic influences, that the natural variability would be less than observed in the 20th century and that, as a consequence, the uncertainty limits estimated by the proposed method are artificial (not representing the natural variability) and too wide. To investigate such eventualities, data of earlier centuries would be needed which are unavailable for the case study area. Besides, the usual technique of splitting the data period into two subperiods and using the early subperiod for model construction and the later subperiod for validation is not applicable in this case because the SSS framework needs long data periods to estimate the Hurst coefficient ($96/2=48$ years would be too few). However, we tried to shed light to these questions using a longer data set, not associated to the above case study in terms of geographical location but closely related in terms of observed statistical behavior. This is the mean annual temperature record of Berlin/Tempelhof, one of the longest series of instrumental meteorological observations going back to 1701. The data set prior to 1993 are made available on the web by the Wetterzentrale/Klimadaten (<http://www.wetterzentrale.de/klima/tberlintem.html>); the most recent data up to 2003 were found from monthly bulletins (Monthly Climatic Data for the World made available on the internet by the National Oceanic and Atmospheric Administration/National Climatic Data Center (<http://www1.ncdc.noaa.gov/pub/data/mcdw/>)). During the period 1756-2003 (248 years) there are no missing data (apart from a few sporadic months which were filled in using the averages for the same month of the closest 4 years). The period 1711-1727 (17 years) and 1752-1754 (3 years) are characterized by complete absence of data. Therefore, all calculations were done based on the continuous period with complete data (1756-2003) but for completeness, approximate climatic values at

the 30-year scale were also calculated for the earlier period by linear interpolation of climate throughout the missing years. The Hurst coefficient for the period 1756-2003 is 0.77; if we split this period into three equal (83-year) subperiods, the Hurst coefficients are 0.83, 0.53 and 0.81 (from the earliest to latest subperiod), which indicates the large uncertainty in the estimation of the Hurst coefficient (this variation is typical of the Hurst behavior; see e.g. Koutsoyiannis, 2003).

The following exercise was done with this data set. The data of the period 1908-2003 (exactly the same with that of the Boeotikos Kephisos case study) were used to estimate the parameters of the scaling model (in this case, $H = 0.78$) and then climatic hindcasts (backward forecasts, prior to 1938 = 1908+30) were calculated in terms of conditional point estimates and confidence intervals (for $\alpha = \alpha' = 95\%$ and climatic time scale of 30 years). This was done for both the scaling and the classical statistical model. The unused part of the series back to 1731 (= 1701+30) was compared with the confidence limits. The results are depicted in Figure 13. For the comparisons of historical values to confidence limits we should have in mind that the confidence coefficient $\alpha' = 95\%$ allows, on the average, one climatic outlier (point outside of the confidence limits) every $20 \times 30 = 600$ years (corresponding to the 30-year time scale); therefore in a single record of about 300 years an outlier is not quite likely; if we also count the parameter uncertainty, which widens confidence intervals, an outlier becomes even less unlikely. However, if the classical model is used, several outliers emerge (the climatic evolution gets out of the confidence limits); even during the model fitting period, the climatic evolution gets out of the unconditional confidence limits. In contrast, the SSS confidence limits capture the historical climatic evolution and simultaneously do not look too wide for the case. This example provides some validation to the method proposed.

5. Conclusion and discussion

A first conclusion of this study is that classical statistics, applied to climatology and hydrology, describes only a portion of natural uncertainty and underestimates seriously the risk if long range dependence, also known as long-term persistence or the Hurst phenomenon, is present. Simple scaling stochastic (SSS) processes offer a sound basis to adapt hydro-

climatic statistics so as to capture interannual variability. The SSS framework, which is developed in this study and is applied to past hydro-climatic records, is a feasible step towards making more reliable estimates of future uncertainty and risk for hydrological studies and water resources management. The uncertainty bands obtained from the SSS framework are about three times wider than those obtained by classical statistics. The detailed case study involving three important hydrometeorological processes (temperature, rainfall and runoff) in a catchment in Greece provides evidence that the SSS, rather than the classical, uncertainty bands are applicable. This was verified further using a much longer instrumental meteorological record (mean annual temperature at Berlin).

However, in this manner the uncertainty estimate is based indirectly on the assumption that past hydroclimatic data are representative for the future, which may not be true if anthropogenic climate changes will be significant compared to natural climatic variability. Therefore, climatic model outputs should be incorporated in an uncertainty analysis and it can be anticipated that future uncertainty is even greater than produced by the SSS framework. In this study, several GCM outputs were examined, which generally agree that temperature will increase significantly, i.e. beyond stationary uncertainty bands, even though different models (rather than different scenarios) predict different levels of temperature increase. At the same time the GCM outputs for rainfall and the resulting (using a hydrological model) runoff do not display significant future changes as the projected time series lie well within SSS confidence limits.

In the case study examined here, a common drawback of the GCM models is the fact that they do not capture past climatic variability, i.e. they result in monthly, annual and over-annual variability that is too low. Obviously, this raises questions for their performance in predicting future climate variability. In this study, a remediation was done to match the historical monthly and annual variability. However, a different procedure should be used for over-annual variability, which as demonstrated is very important for the determination of future uncertainty. This should be the subject of future research. The general idea of such a research would be the representation of the variability of processes in stochastic terms using the historical data and the representation of changes in the mean of the processes based on

GCM results. In other words, GCM would be used to establish causative relationships for mean levels of processes. The establishment of such causative relationships would probably require steady-state GCM runs, with stable forcing conditions (e.g. different CO₂ and SO₄ concentrations) rather than varying scenario-based conditions.

Appendix: Derivation of the equations for the conditional climatic expectation and its variance

Using a known result (Koutsoyiannis, 2000), it is easily shown that the conditional mean of a process X_i at the scale 1 is given as

$$E[X_i|\mathbf{x}_{0,n}] = \mu + \boldsymbol{\rho}'_{i,n} \boldsymbol{\tau}_n^{-1} (\mathbf{x}_{0,n} - \mathbf{e}_n \mu) = \mu(1 - \boldsymbol{\rho}'_{i,n} \boldsymbol{\tau}_n^{-1} \mathbf{e}_n) + \boldsymbol{\rho}'_{i,n} \boldsymbol{\tau}_n^{-1} \mathbf{x}_{0,n} \quad (\text{A.1})$$

where the notation is explained in section 3.4. Consequently, at scale i ,

$$E[X_i^{(i)}|\mathbf{x}_{0,n}] = \mu(1 - \boldsymbol{\xi}'_{i,n} \boldsymbol{\tau}_n^{-1} \mathbf{e}_n) + \boldsymbol{\xi}'_{i,n} \boldsymbol{\tau}_n^{-1} \mathbf{x}_{0,n} \quad (\text{A.2})$$

Thus,

$$\begin{aligned} E[X_i^{(k)}|\mathbf{x}_{0,n}] &= \left[\frac{i}{k} (1 - \boldsymbol{\xi}'_{i,n} \boldsymbol{\tau}_n^{-1} \mathbf{e}_n) + \left(1 - \frac{i}{k}\right) (1 - \boldsymbol{\xi}'_{i-k,n} \boldsymbol{\tau}_n^{-1} \mathbf{e}_n) \right] \mu \\ &\quad + \left[\frac{i}{k} \boldsymbol{\xi}'_{i,n} + \left(1 - \frac{i}{k}\right) \boldsymbol{\xi}'_{i-k,n} \right] \boldsymbol{\tau}_n^{-1} \mathbf{x}_{0,n}, \quad i \geq k \end{aligned} \quad (\text{A.3})$$

$$E[X_i^{(k)}|\mathbf{x}_{0,n}] = \frac{i}{k} (1 - \boldsymbol{\xi}'_{i,n} \boldsymbol{\tau}_n^{-1} \mathbf{e}_n) \mu + \frac{i}{k} \boldsymbol{\xi}'_{i,n} \boldsymbol{\tau}_n^{-1} \mathbf{x}_{0,n} + \left(1 - \frac{i}{k}\right) x_0^{(k-i)}, \quad i \leq k \quad (\text{A.4})$$

Recall that $1 - \boldsymbol{\xi}'_{i,n} \boldsymbol{\tau}_n^{-1} \mathbf{e}_n =: \varphi_{i,n}(H)$ is a function of the Hurst coefficient H , a lag i and the observation length n . For fixed $n = 100$, which is close to the sample sizes used in this study the function $\varphi_{i,n}(H)$ (shown in Figure 7, left panel), is almost perfectly approximated by the following equations, which were derived by an extended numerical investigation:

$$\varphi_{i,100}(H) = 1 - (2H - 1)^{c_1} [1 - c_2 (1 - H)] \quad (\text{A.5})$$

where

$$c_1 = 0.75 + 0.1 \ln i, \quad c_2 = 2 - 3.3 \exp[-(0.18 \ln i)^{3.7}] \quad (\text{A.6})$$

It can be observed that: (a) for $H = 0.5$, $\varphi = 1$ as expected (no effect from the past which is consistent with independence); (b) for $i \rightarrow \infty$, $c_1 \rightarrow \infty$ and since $(2H - 1) < 1$, again $\varphi = 1$ (no

effect from the past because of large lag). However, the convergence of φ to 1 is very slow, unless H is close to 0.5.

It is observed from (A.3)-(A.4), which are exact equations, that the conditional mean in addition to μ and φ , depends, on a linear combination of the entire observation vector $\mathbf{x}_{0,n}$. To make the expression simpler (which is important for the Monte Carlo simulation) and more insightful (thus allowing the explanations given in section 3.4), we can set $\mathbf{x}_{0,n} \approx x_0^{(n)} \mathbf{e}$ (where $x_0^{(n)}$ is the sample mean). In this manner, we readily obtain (18) and (19).

For the conditional variance of the climatic process, we start from a known result (Koutsoyiannis, 2000) that the conditional variance of a variable X_i at the scale 1 is given as

$$\text{Var}[X_i|\mathbf{x}_{0,n}] = \sigma^2(1 - \boldsymbol{\rho}'_{i,n} \boldsymbol{\tau}_n^{-1} \boldsymbol{\rho}_{i,n}) \quad (\text{A.7})$$

Recall that $1 - \boldsymbol{\rho}'_{i,n} \boldsymbol{\tau}_n^{-1} \boldsymbol{\rho}_{i,n} =: \psi_{i,n}(H)$ is a function of the Hurst coefficient H , a lag i and the observation length n . As $n \rightarrow \infty$, it tends to a function $\psi_i(H)$ (plotted in Figure 7, right panel); numerical investigation shows that the convergence is fast, so that for n as low as 5 to 10, $\psi_i(H)$ is a good approximation for $\psi_{i,n}(H)$. Therefore, in all analyses we use the asymptotic function $\psi_i(H)$ only. By a further numerical investigation, we were able to obtain the following approximate expression of $\psi_i(H)$, which is almost perfect

$$\psi_i(H) = 1 - (2H - 1)^{2 + \ln i} [1 - (2 - 1.28 / i^{0.25}) (1 - H)] \quad (\text{A.8})$$

For $H = 0.5$ it can be directly verified that $\psi_i(H) = 1$ (no attenuation, because of temporal independence). Furthermore, for $i \rightarrow \infty$, it can be again verified that $\psi_i(H) \rightarrow 1$ (no attenuation because of large lag, despite the high temporal dependence; however, the convergence to 1 is very slow, unless H is close to 0.5). Finally, for $i = 1$ the obtained $\psi_1(H)$ is identical to an expression derived in Koutsoyiannis (2005b).

In the SSS model the autocorrelation function is invariant to the reference time scale (e.g. Koutsoyiannis, 2003). Thus, the function $\psi_i(H)$ does not depend on scale; consequently for a scale k and a lag i assumed to be an integer multiple of k , i.e. $i = l k$, where $l = 1, 2, \dots$, from (A.7) we obtain

$$\text{Var}[X_{l,k}^{(k)} | \mathbf{x}_{0,\infty}^{(k)}] = [\sigma^{(k)}]^2 \psi_l(H) \quad (\text{A.9})$$

where $\mathbf{x}_{0,\infty}^{(k)}$ is the vector of realization of the present and the entire past. Obviously, the entire past $\mathbf{x}_{0,\infty}^{(k)}$ cannot be known and only $\mathbf{x}_{0,n}$ is observed. However, given the already mentioned fast convergence of $\psi_{l,n}(H)$ to $\psi_l(H)$ and assuming that the time scale of observation of the past does not significantly influence the information for predictions in the future, we may substitute the condition $\mathbf{x}_{0,n}$ for $\mathbf{x}_{0,\infty}^{(k)}$ to obtain an approximation of the conditional variance for the actually observed past. If we also substitute $\sigma^{(k)}$ from (6) we obtain the approximation

$$\text{StD}[X_{l,k}^{(k)} | \mathbf{x}_{0,n}] = k^{H-1} \sigma \sqrt{\psi_l(H)} \quad (\text{A.10})$$

For $l \geq 1$ (or $i = l$ $k \geq k$), (A.10) gives directly the required standard deviation of the climatic variable $X_i^{(k)}$ conditional on the observations of the past. For $i \leq k$, by virtue of (17) we obtain from (A.10)

$$\text{StD}[X_i^{(k)} | \mathbf{x}_{0,n}] = \frac{i}{k} \text{StD}[X_i^{(i)} | \mathbf{x}_{0,n}] = \frac{i^H}{k} \sigma \sqrt{\psi_1(H)} \quad (\text{A.11})$$

From (A.10) and (A.11) we readily obtain (22) and (23), respectively.

Acknowledgments

Dr. Jianzhong Wang of the Hydrologic Research Center (HRC) assisted with the preparation of GCM data. Numerous comments from four anonymous reviewers resulted in improved presentation of the ideas. DK wishes to thank KG and the HRC staff for their warm hospitality during his visit.

References

- Anderson, J.L., van den Dool, H., Barnston, A., Chen, W., Stern, W., and J. Ploshay, 1999: Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bull. Amer. Meteor. Soc.* **80**, 1349-1361.
- Beran, J., 1994: *Statistics for Long-Memory Processes*, Volume 61 of *Monographs on Statistics and Applied Probability*, Chapman and Hall, New York.
- Bickel, P., and Doksum, K., 2000: *Mathematical Statistics: Basic Ideas and Selected Topics*, volume I, second edition, Prentice Hall.
- Bhattacharya, R. N., Gupta, V. K. and Waymire, E., 1983: The Hurst effect under trends, *J. Appl. Prob.*, 20, 649–662.
- Carter, T.R., M. Hulme and M. Lal, 1999: Guidelines on the Use of Scenario Data for Climate Impact and Adaptation Assessment, Task Group on Scenarios for Climate Impact Assessment, Intergovernmental Panel on Climate Change (http://ipcc-ddc.cru.uea.ac.uk/guidelines/ggm_no1_v1_12-1999.pdf).
- Cayan, D.R., Kammerdiener, S.A., Dettinger, M.D., Caprior, J.M., and D.H. Peterson, 2001: Changes in the onset of spring in the western US. *Bull. Am. Met. Soc.* **82**, 399-415,.
- Cohn, T., and H. Lins, 2005: Nature's style: Naturally trendy, *Geophysical Research Letters*, 32(23), art. no. L23402 .
- Dettinger, M.D., Cayan, D.R., Meyer, M.K., and A.E. Jeton, 2004: Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American river basins, Sierra Nevada, California, 1900-2099. *Clim. Change* **62**, 283-317.
- Efstratiadis, A., E. Rozos, A. Koukouvinos, I. Nalbantis, G. Karavokiros, and D. Koutsoyiannis, 2005: An integrated model for conjunctive simulation of hydrological processes and water resources management in river basins, *2nd General Assembly of the European Geosciences Union, Geophysical Research Abstracts*, Vol. 7, Vienna, 03560, European Geosciences Union.

- Flato, G.M. and Boer, G. J., 2001: Warming asymmetry in climate change simulations, *Geophys. Res. Lett.*, 28, 195-198.
- Georgakakos, K.P., 2003: Probabilistic Climate-Model Diagnostics for Hydrologic and Water Resources Impact Studies, *Journal of Hydrometeorology*, 4, 92-105.
- Georgakakos, K.P. and D.E. Smith, 2001: Soil moisture tendencies into the next century for the conterminous United States, *Journal of Geophysical Research - Atmosphere*, 106(D21), 27367-27382.
- Giorgi, F., Whetton, P.H., and R.G. Jones, 2001: Emerging patterns of simulated regional climatic changes for the 21st century due to anthropogenic forcings. *Geoph. Res. Letters* 28, 3317-3321.
- Goddard, L., Mason, S.J., Zebiak, S.E., Ropelewski, C.F., Basher, R., and M.A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. Journal of Climatology*, 21, 1111-1152.
- Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A., 2000: The simulation of SST, sea ice extents and ocean heat transports in aversion of the Hadley Centre Coupled Model without flux adjustments, *Climate Dynamics* 16 147-168
- Hirsch R. M., Helsel, D. R., Cohn, T. A., and Gilroy, E. J., 1993: Statistical analysis of hydrological data, in *Handbook of Hydrology*, D. Maidment (ed.), McGraw-Hill.
- Hurst, H. E., 1951: Long term storage capacities of reservoirs, *Trans. ASCE*, 116, 776-808.
- International Panel for Climate Change (IPCC) 2001: Climate change 2001, Impacts, Adaptation and Vulnerability (eds. McCarthy, J.J., et al.). Cambridge University press, Cambridge, UK.
- Kantelhardt, J. W, S. A. Zschiegnera, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, 2002: Multifractal detrended fluctuation analysis of nonstationary time series, *Physica A*, 316, 87 – 114.
- Kharin, V.V., and F.W. Zwiers, 2000: Changes in the Extremes in an Ensemble of Transient Climate Simulations with a Coupled Atmosphere-Ocean GCM. *Journal of Climate*, 13, 3760-3788.

- Klemeš, V., 1974: The Hurst phenomenon: a puzzle?, *Wat. Resour. Res.* 10(4), 675–688.
- Klemeš, V., 1978: Physically based stochastic hydrologic analysis, *Advances in Hydroscience*, 11, 285-356.
- Koutsoyiannis, D., 2000: A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research*, 36(6), 1519-1534.
- Koutsoyiannis, D., 2002: The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrologic Sciences Journal*, 47(4), 573-595.
- Koutsoyiannis, D., 2003: Climate change, the Hurst phenomenon, and hydrologic statistics, *Hydrologic Sciences Journal*, 48(1), 3-24.
- Koutsoyiannis, D., 2005a: Uncertainty, entropy, scaling and hydrologic stochastics, 1, Marginal distributional properties of hydrological processes and state scaling, *Hydrologic Sciences Journal*, 50(3), 381-404.
- Koutsoyiannis, D., 2005b: Uncertainty, entropy, scaling and hydrologic stochastics, 2, Time dependence of hydrologic processes and time scaling, *Hydrologic Sciences Journal*, 50(3), 405-426.
- Koutsoyiannis, D., 2006a: A toy model of climatic variability with scaling behaviour, *Journal of Hydrology* 322, 25-48, 2006.
- Koutsoyiannis, D., 2006b: Nonstationarity versus scaling in hydrology, *Journal of Hydrology* 324, 239-254.
- Koutsoyiannis, A., Efstratiadis, A. and Georgakakos, K. P., 2006: Uncertainty assessment of future hydroclimatic predictions: A comparison of probabilistic and scenario-based approaches - Additional information, *Internal report*, Department of Water Resources, National Technical University of Athens (<http://www.itia.ntua.gr/e/docinfo/728/>).
- Koutsoyiannis, D. and S. Kozanis, 2005: A simple Monte Carlo methodology to calculate generalized approximate confidence intervals, HRC Technical Note No. 25, Hydrologic Research Center, San Diego (<http://www.itia.ntua.gr/e/docinfo/692/>).
- Koutsoyiannis, D., and Th. Xanthopoulos, 1999: *Engineering Hydrology* (in Greek), Edition 3, 418 pages, National Technical University of Athens, Athens.

- Leggett, J., W.J. Pepper and R.J. Swart, 1992: Emissions scenarios for the IPCC: an update, in: *Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment*, edited by Houghton, J.T., B.A. Callander, and S.K. Varney, Cambridge University Press, Cambridge, pp. 75-95.
- Mason, S.J., Goddard, L., Graham, N.E., Yulaeva, E., Sun, L., and P.A. Arkin, 1999: The IRI Seasonal Climate Prediction System and the 1997/98 El Niño Event. *Bull. Amer. Meteor. Soc.* **80**, 1853-1873.
- Mesa, O. J. and Poveda, G., 1993: The Hurst effect: the scale of fluctuation approach, *Wat. Resour. Res.* 29(12), 3995–4002.
- Montanari, A., Taqqu, M.S. and Teverovsky, V., 1999: Estimating long-range dependence in the presence of periodicity: an empirical study. *Mathematical and Computer Modeling*, 29, 217-228,.
- Nakicenovic, N. and R. Swart (eds), 1999: IPCC Special Report on Emissions Scenarios, Intergovernmental Panel on Climate Change (<http://www.grida.no/climate/ipcc/emission/>).
- Papoulis, A., 1990: *Probability and Statistics*, Prentice-Hall, New Jersey.
- Risbey, J.S. and P.H. Stone, 1996: A case study of the adequacy of GCM simulations for input to regional climate change assessments. *J. Climate*, **9**, 1441-1467.
- Rozos, E., A. Efstratiadis, I. Nalbantis, and D. Koutsoyiannis, 2004: Calibration of a semi-distributed model for conjunctive simulation of surface and groundwater flows, *Hydrologic Sciences Journal*, 49(5), 819-842.
- Shukla, J., Anderson, D. Baumhefner, C. Brankovic, Y. Chang, E. Kalnay, L. Marx, T. Palmer, D. Paolino, J. Ploshay, S. Schubert, D. Straus, M. Suarez, and J. Tribbia, 2000: Dynamical Seasonal Prediction. *Bull. Amer. Meteor. Soc.* **81**, 2593-2606.
- Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou, 1993: Frequency analysis of extreme events, Chapter 18 in *Handbook of Hydrology*, edited by D. R. Maidment, McGraw-Hill.
- Straus, D., 1993: The Midlatitude Development of Regional Errors in a Global GCM. *Journal of the Atmospheric Sciences*, **50**(16), 2785-2799.

Tsonis, A. A., 1992: *Chaos: From Theory to Applications*, 274 pp., Plenum, New York.

Wang, X.L., and F.W. Zwiers, 1999: Interannual Variability of Precipitation in an Ensemble of AMIP Climate Simulations Conducted with the CCC GCM2. *Journal of Climate*, **12**, 1322-1335.

Figure captions

Figure 1 The Boeoticos Kephisos River basin (left) and location key (right), in which triangles indicate the grid points of GCMs nearest to the catchment; their coordinates are: (37.6731°N, 22.50°E) for M1 (MP01GG01 and MP01GS01); (38.9666°N, 22.50°E) for M2 (CCCma_A2 and B2); and (37.50°N, 22.50°E) for M3 (HADCM3_A2 and B2).

Figure 2 Explanation sketch for the total uncertainty of a climatic variable.

Figure 3 Logarithmic plots of standard deviation versus scale for the time series of (upper left) Aliartos temperature (°C) (upper right) Aliartos rainfall (mm) (lower left) Boeoticos Kephisos annual runoff (mm) ; (lower right) a synthetic time series with length and statistical characteristics equal to those of Boeoticos Kephisos annual runoff but assuming independence.

Figure 4 Return periods of the minimum and maximum values of the average runoff, over scale $k = 1$ to 10 years, for the Boeoticos Kephisos 96-year runoff record; the return periods were calculated assuming that the distribution is normal for all scales and that the standard deviation over scale k is given by (5) and (6) for the classical and SSS model, respectively.

Figure 5 Point estimates and confidence limits (for $\alpha = 95\%$) of distribution quantiles of (upper) annual and (lower) climatic (30-year) runoff of Boeoticos Kephisos. (PE: point estimate; MCCL/classical: Monte Carlo confidence limits assuming independence as in classical statistics; MCPL/SSS 1: Monte Carlo confidence limits assuming scaling and known H ; MCPL/SSS 2: : Monte Carlo confidence limits assuming scaling and unknown H).

Figure 6 Point estimates and confidence limits (for $\alpha = 95\%$) of distribution quantiles of the climatic (30-year) values of (upper) rainfall at Aliartos and (lower) runoff of Boeoticos Kephisos (see Figure 5 for explanation of abbreviations).

Figure 7 Functions $\varphi_{i,n}(H)$ and $\psi_f(H)$ involved in the estimation of conditional climatic confidence limits (for definitions see text).

Figure 8 Annual and climatic means conditional on observations for the Boeoticos Kephisos runoff. Year 0 corresponds to hydrologic year 2002-03. The true parameters were assumed equal to their sample estimates. Exact quantities are estimated from (A.4) and approximate quantities are estimated from (19).

Figure 9 Ratios of conditional mean and standard deviations to the unconditional quantities. The true parameters were assumed equal to their sample estimates. Three additive components are given for the mean corresponding to the influences of the true mean, the sample mean and the constant term, as derived in equation (19).

Figure 10 Historical climate and (conditional) point estimates and confidence limits of future climate (for $\alpha = \alpha' = 95\%$ and climatic time scale of 30 years) for (upper) temperature at Aliartos, (middle) rainfall at Aliartos and (lower) runoff of Boeoticos Kephisos.

Figure 11 Comparison of GCM modeled and historical temperature (upper) and rainfall (lower) for period 1960-89 in terms of their monthly means and standard deviations.

Figure 12 GCM predictions of past and future evolution of hydroclimatic processes at the Boeoticos Kephisos river basin vs. the uncertainty limits; “Historical 1” is measured data whereas “Historical 2” in the lower panel was produced by the hydrological model with historical rainfall and temperature inputs.

Figure 13 Historical climatic temperature at Berlin/Tempelhof and stochastic model hindcasts before 1938, based on the annual data of the period 1908-2003. The hindcasts are in terms of conditional point estimates and confidence limits (for $\alpha = \alpha' = 95\%$ and climatic time scale of 30 years).



Figure 1 The Boeoticos Kephisos River basin (left) and location key (right), in which triangles indicate the grid points of GCMs nearest to the catchment; their coordinates are: (37.6731°N, 22.50°E) for M1 (MP01GG01 and MP01GS01); (38.9666°N, 22.50°E) for M2 (CCCma_A2 and B2); and (37.50°N, 22.50°E) for M3 (HADCM3_A2 and B2).

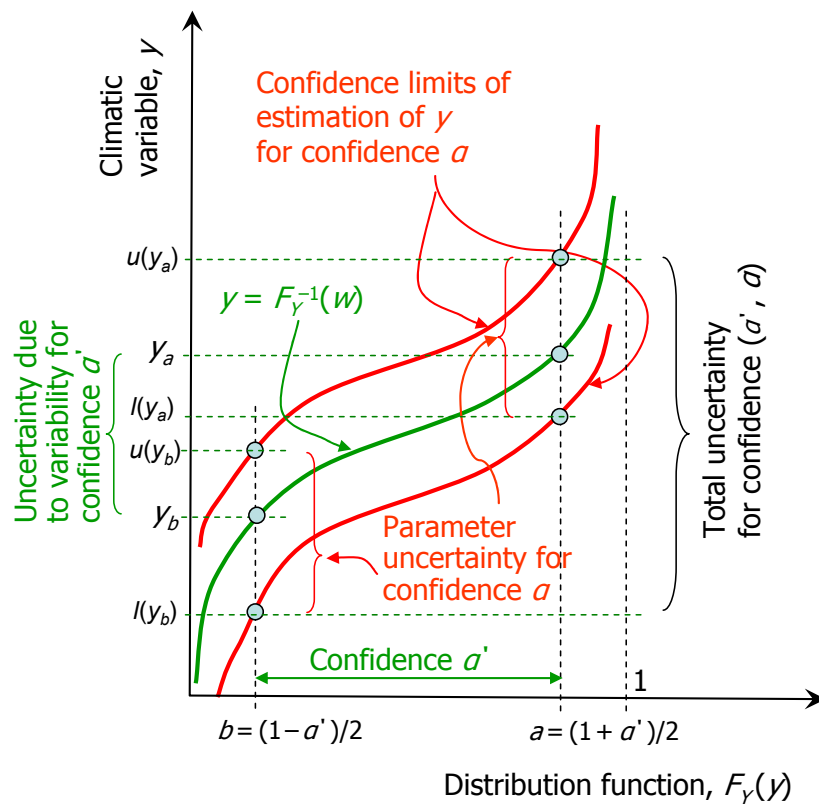


Figure 2 Explanation sketch for the total uncertainty of a climatic variable.

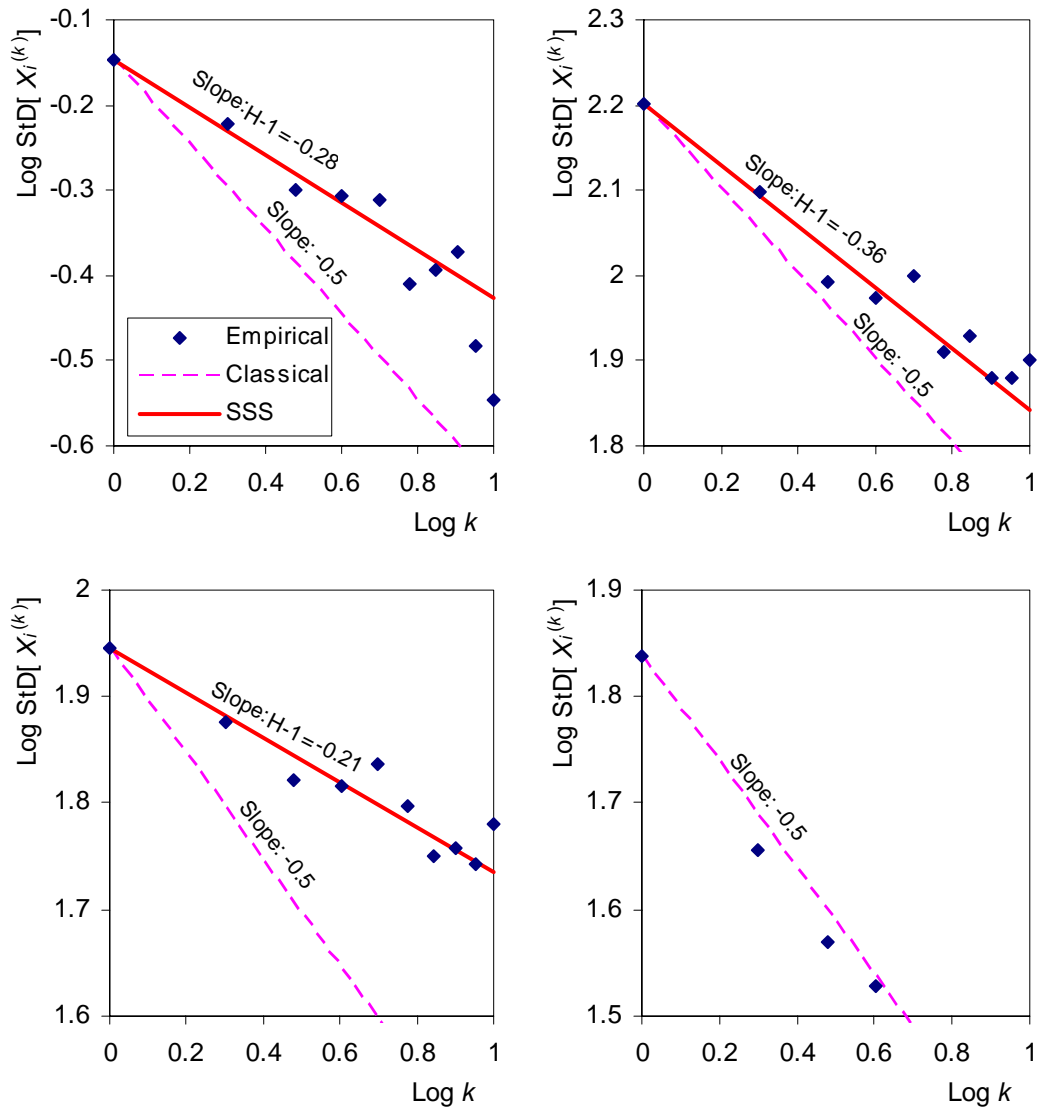


Figure 3 Logarithmic plots of standard deviation versus scale for the time series of (upper left) Aliartos temperature ($^{\circ}\text{C}$) (upper right) Aliartos rainfall (mm) (lower left) Boeoticos Kephisos annual runoff (mm) ; (lower right) a synthetic time series with length and statistical characteristics equal to those of Boeoticos Kephisos annual runoff but assuming independence.

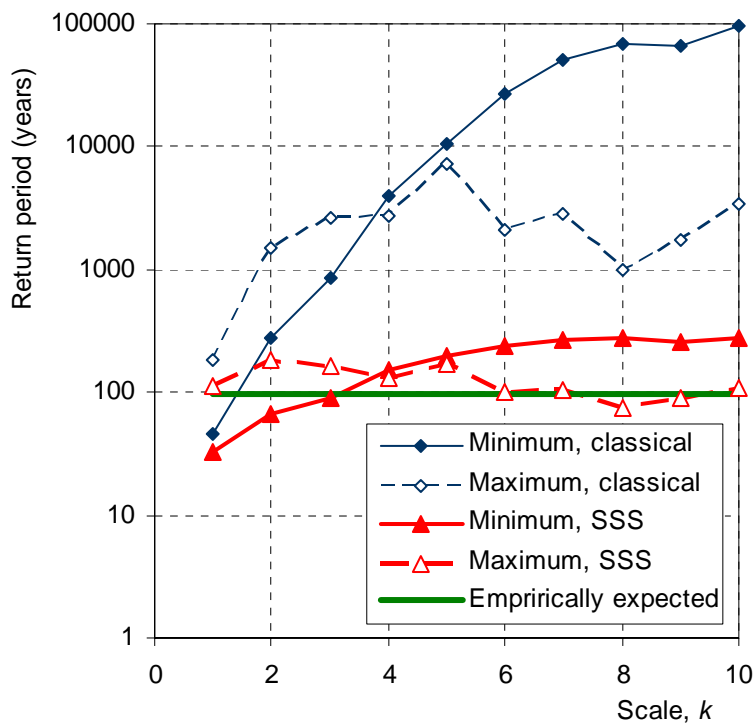


Figure 4 Return periods of the minimum and maximum values of the average runoff, over scale $k = 1$ to 10 years, for the Boeotikos Kephisos 96-year runoff record; the return periods were calculated assuming that the distribution is normal for all scales and that the standard deviation over scale k is given by (5) and (6) for the classical and SSS model, respectively.

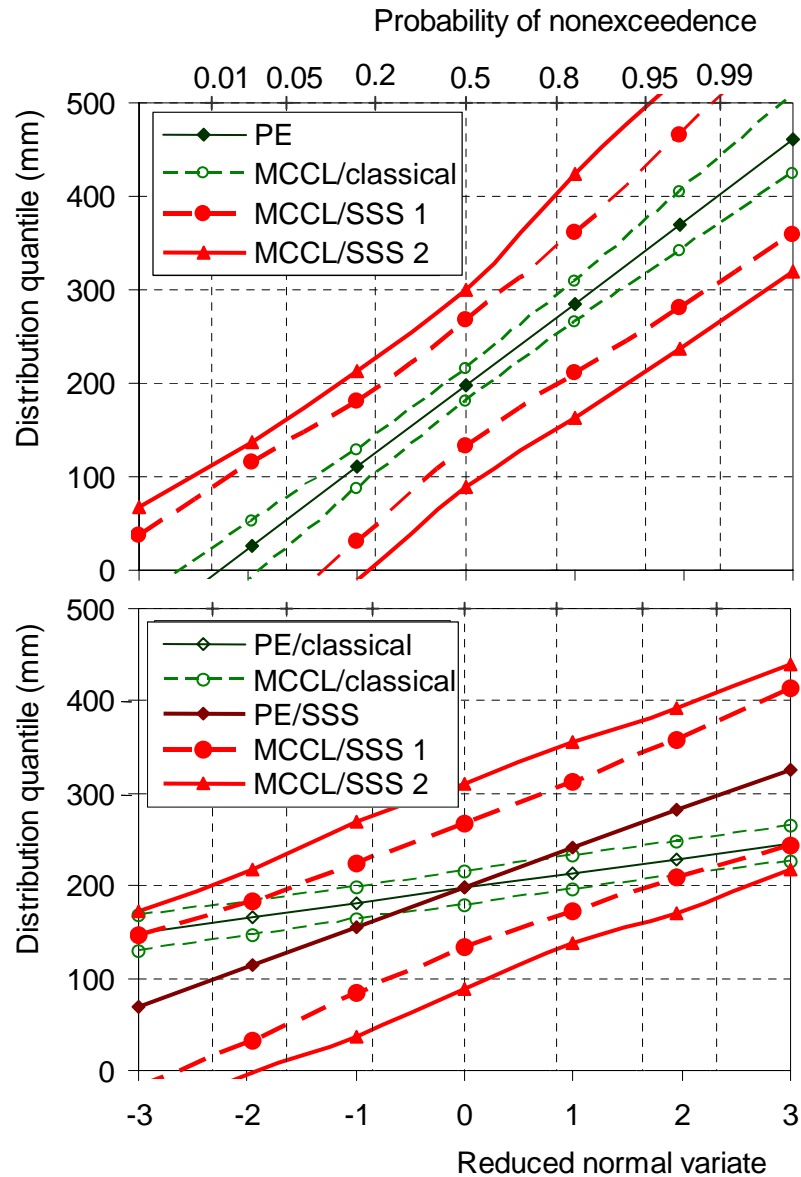


Figure 5 Point estimates and confidence limits (for $\alpha = 95\%$) of distribution quantiles of (upper) annual and (lower) climatic (30-year) runoff of Boeoticos Kephisos. (PE: point estimate; MCCL/classical: Monte Carlo confidence limits assuming independence as in classical statistics; MCPL/SSS 1: Monte Carlo confidence limits assuming scaling and known H ; MCPL/SSS 2: : Monte Carlo confidence limits assuming scaling and unknown H).

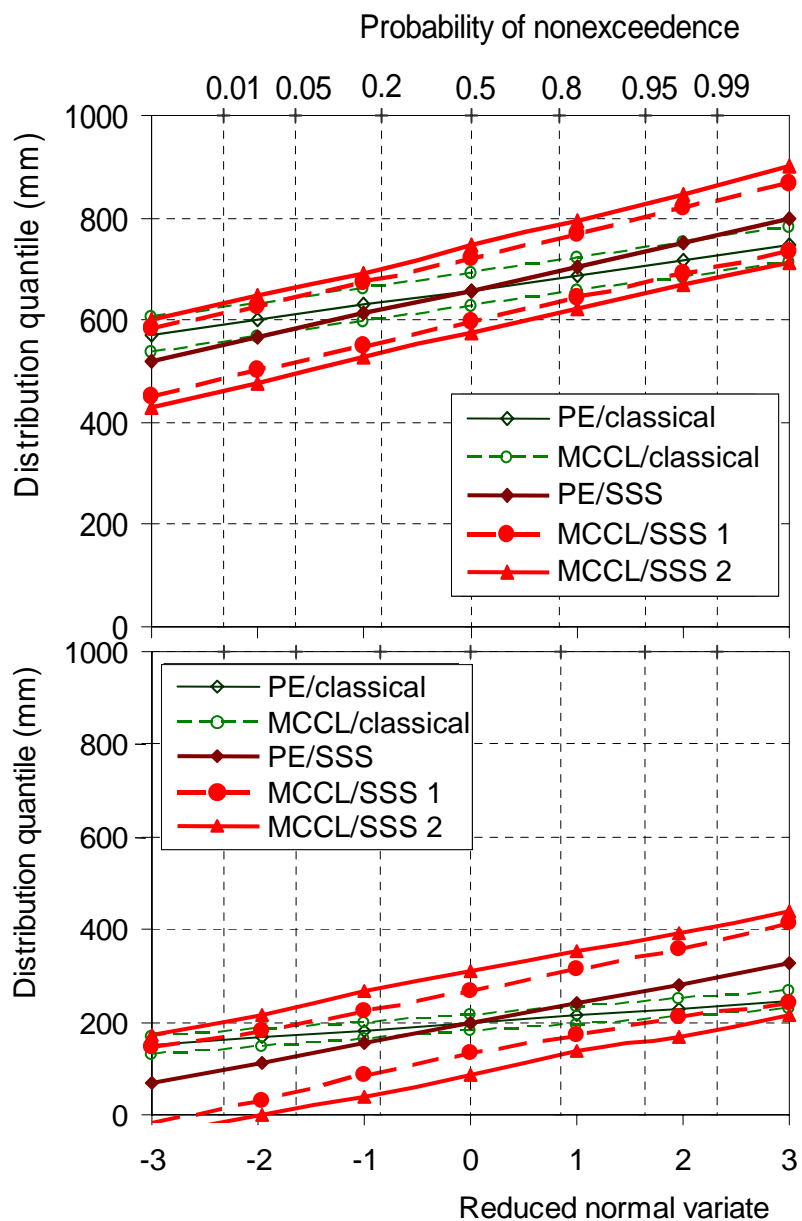


Figure 6 Point estimates and confidence limits (for $\alpha = 95\%$) of distribution quantiles of the climatic (30-year) values of (upper) rainfall at Aliartos and (lower) runoff of Boeticos Kephisos (see Figure 5 for explanation of abbreviations).

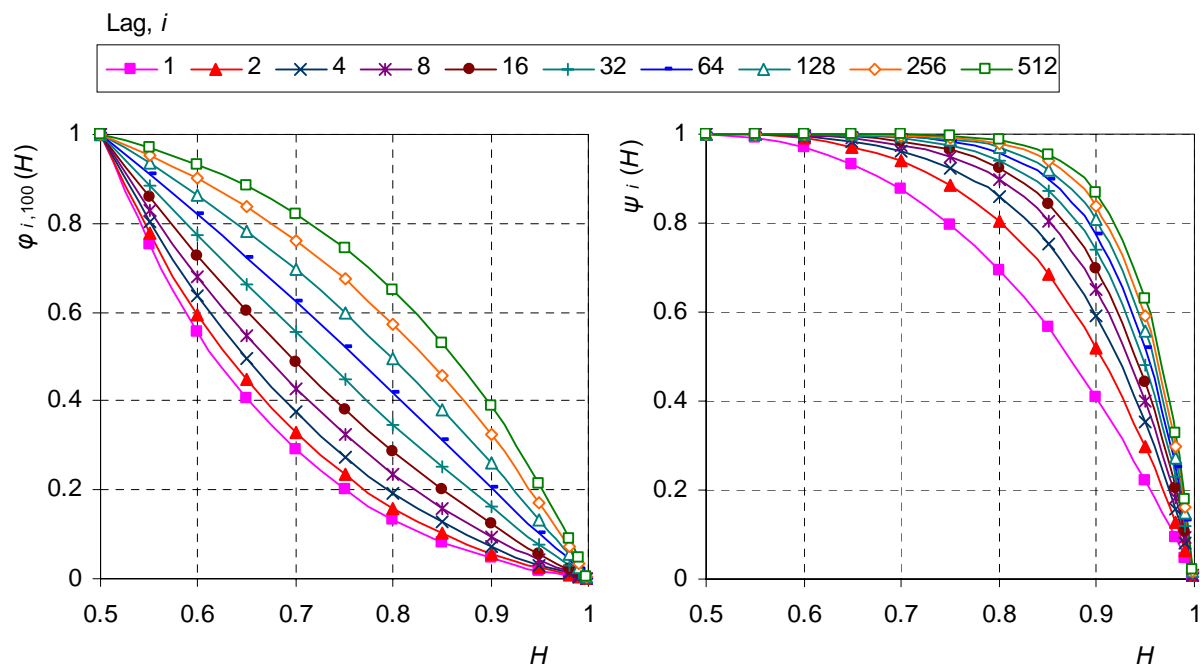


Figure 7 Functions $\varphi_{i,n}(H)$ and $\psi_i(H)$ involved in the estimation of conditional climatic confidence limits (for definitions see text).

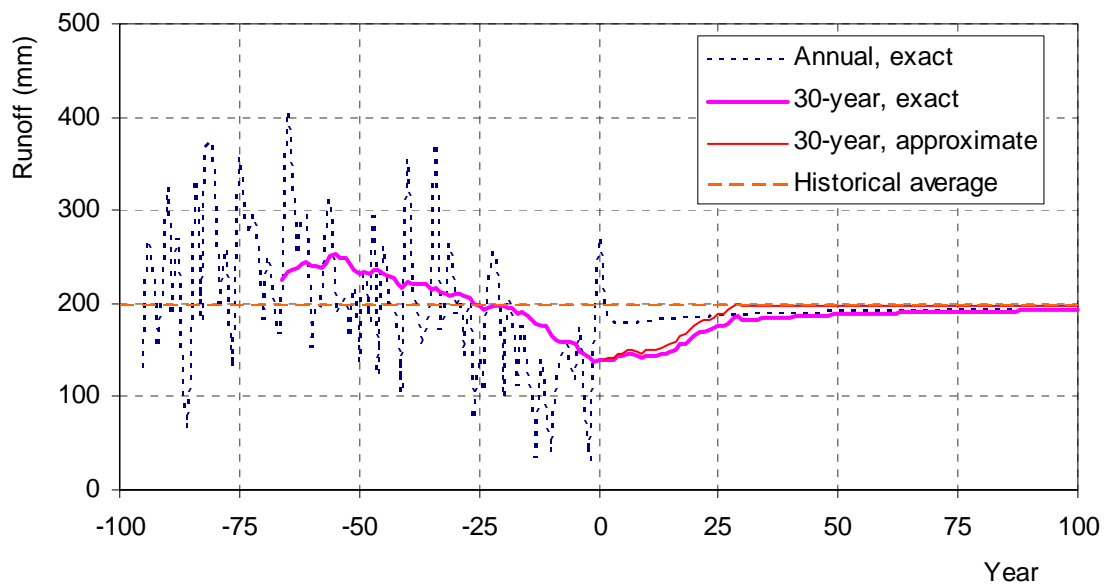


Figure 8 Annual and climatic means conditional on observations for the Boeoticos Kephisos runoff. Year 0 corresponds to hydrologic year 2002-03. The true parameters were assumed equal to their sample estimates. Exact quantities are estimated from (A.4) and approximate quantities are estimated from (19).

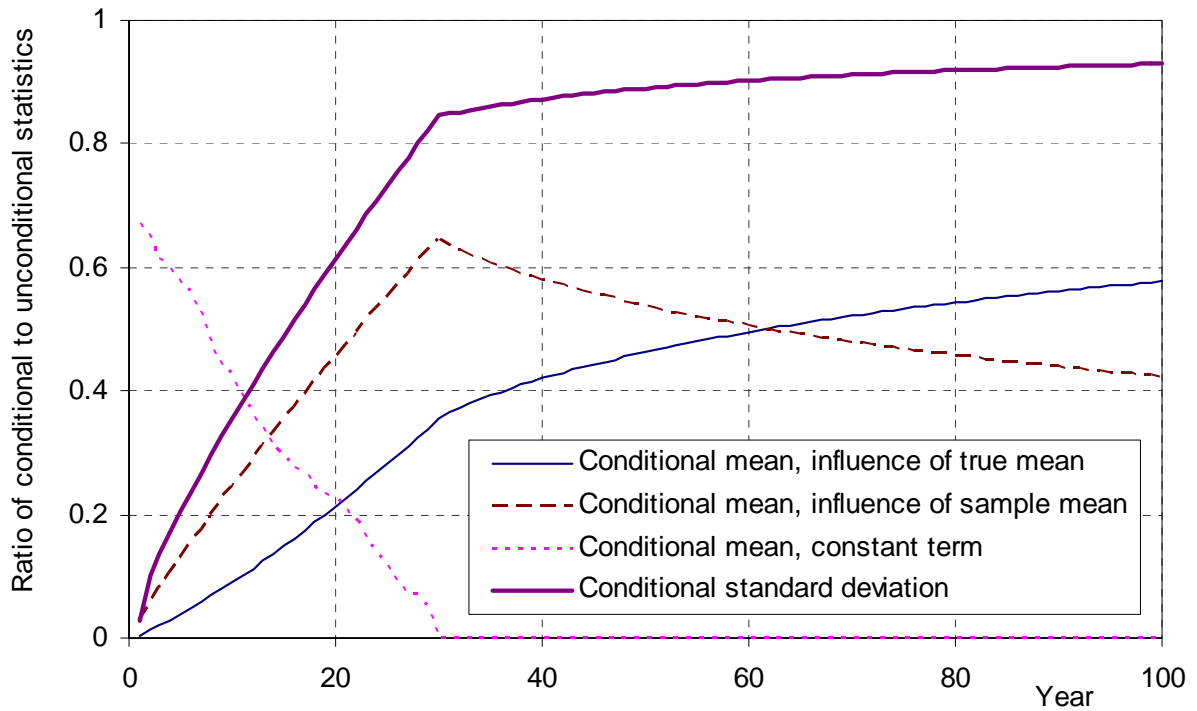


Figure 9 Ratios of conditional mean and standard deviations to the unconditional quantities. The true parameters were assumed equal to their sample estimates. Three additive components are given for the mean corresponding to the influences of the true mean, the sample mean and the constant term, as derived in equation (19).

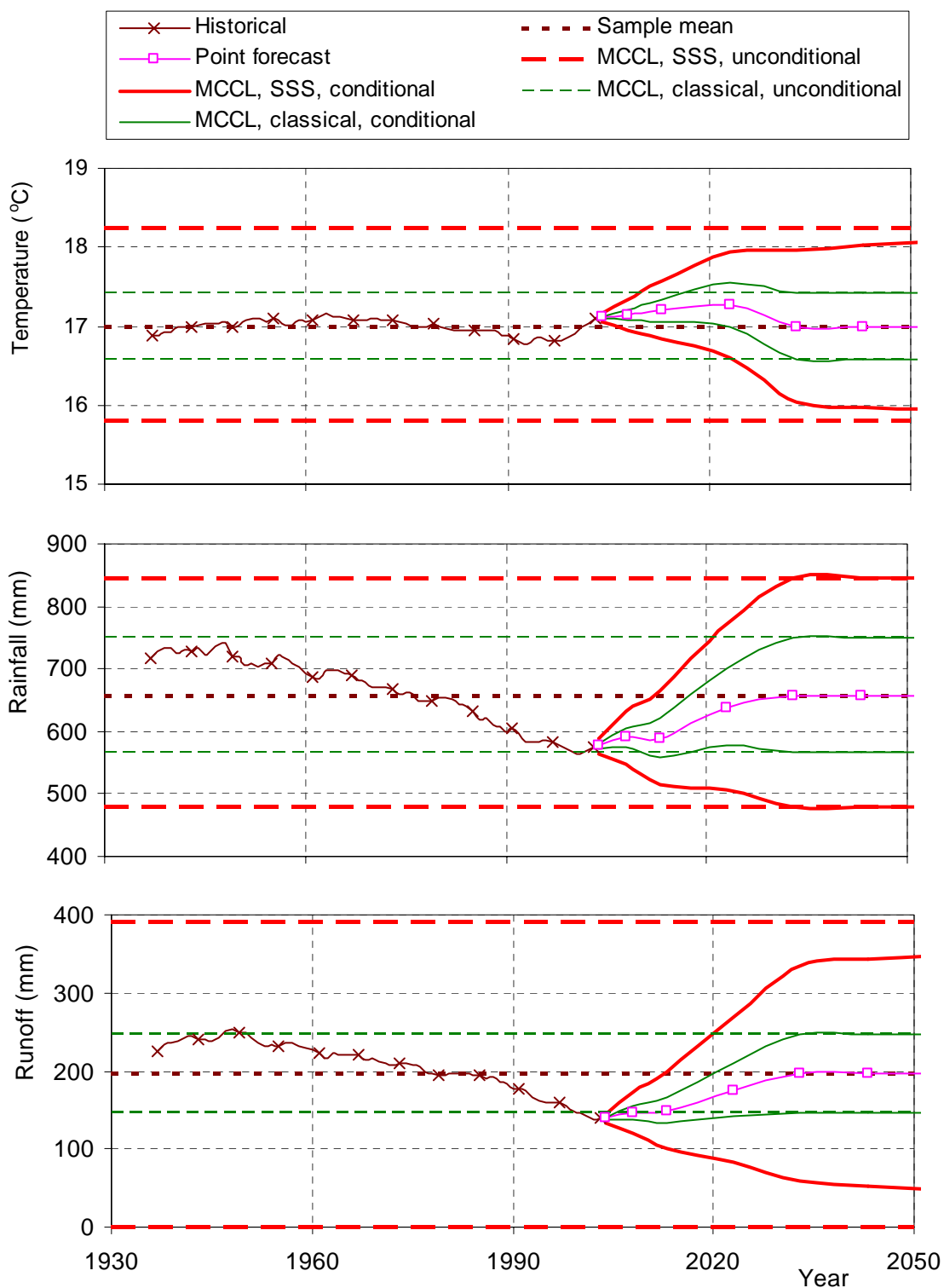


Figure 10 Historical climate and (conditional) point estimates and confidence limits of future climate (for $\alpha = \alpha' = 95\%$ and climatic time scale of 30 years) for (upper) temperature at Aliartos, (middle) rainfall at Aliartos and (lower) runoff of Boeoticos Kephisos.

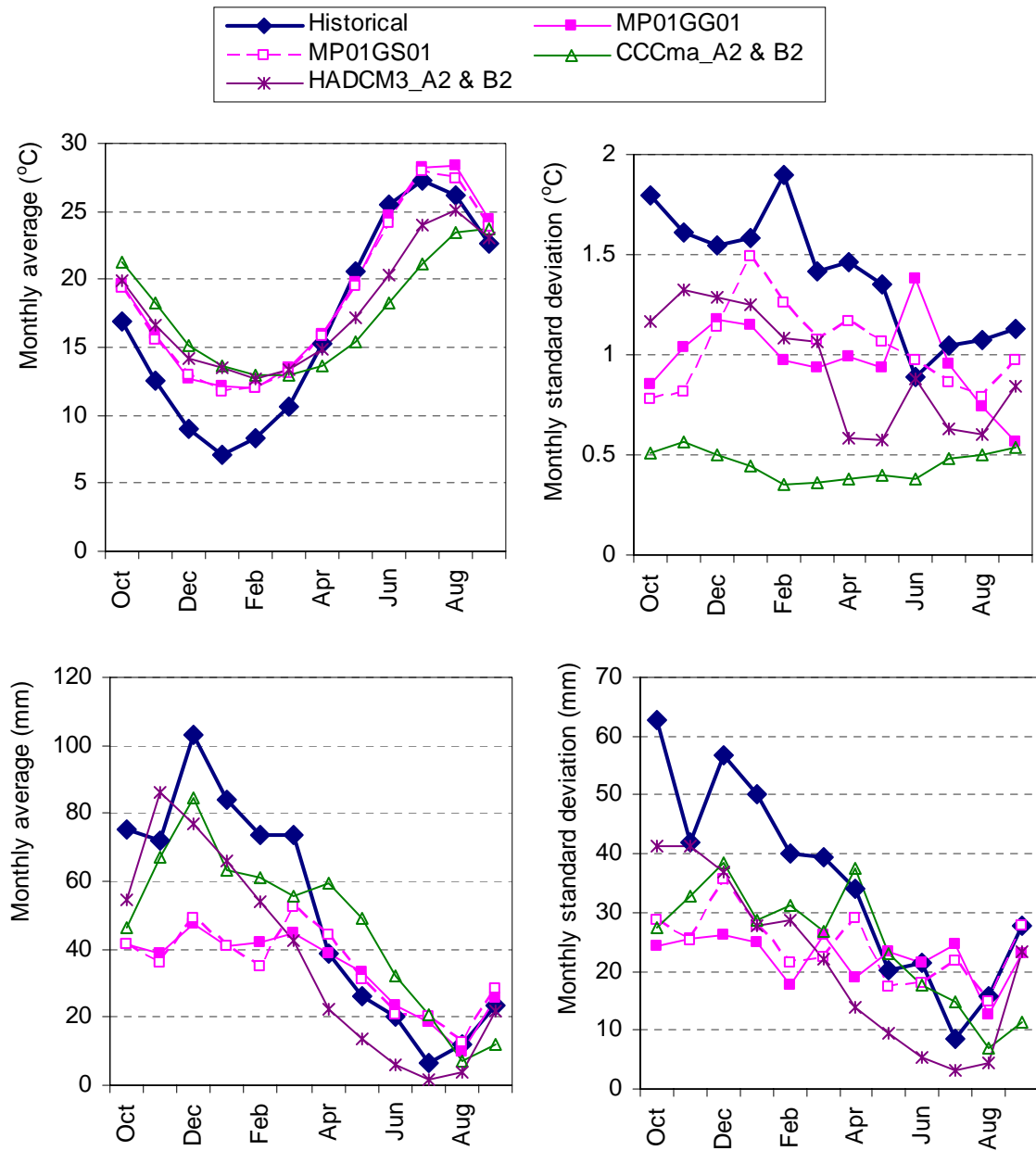


Figure 11 Comparison of GCM modeled and historical temperature (upper) and rainfall (lower) for period 1960-89 in terms of their monthly means and standard deviations.

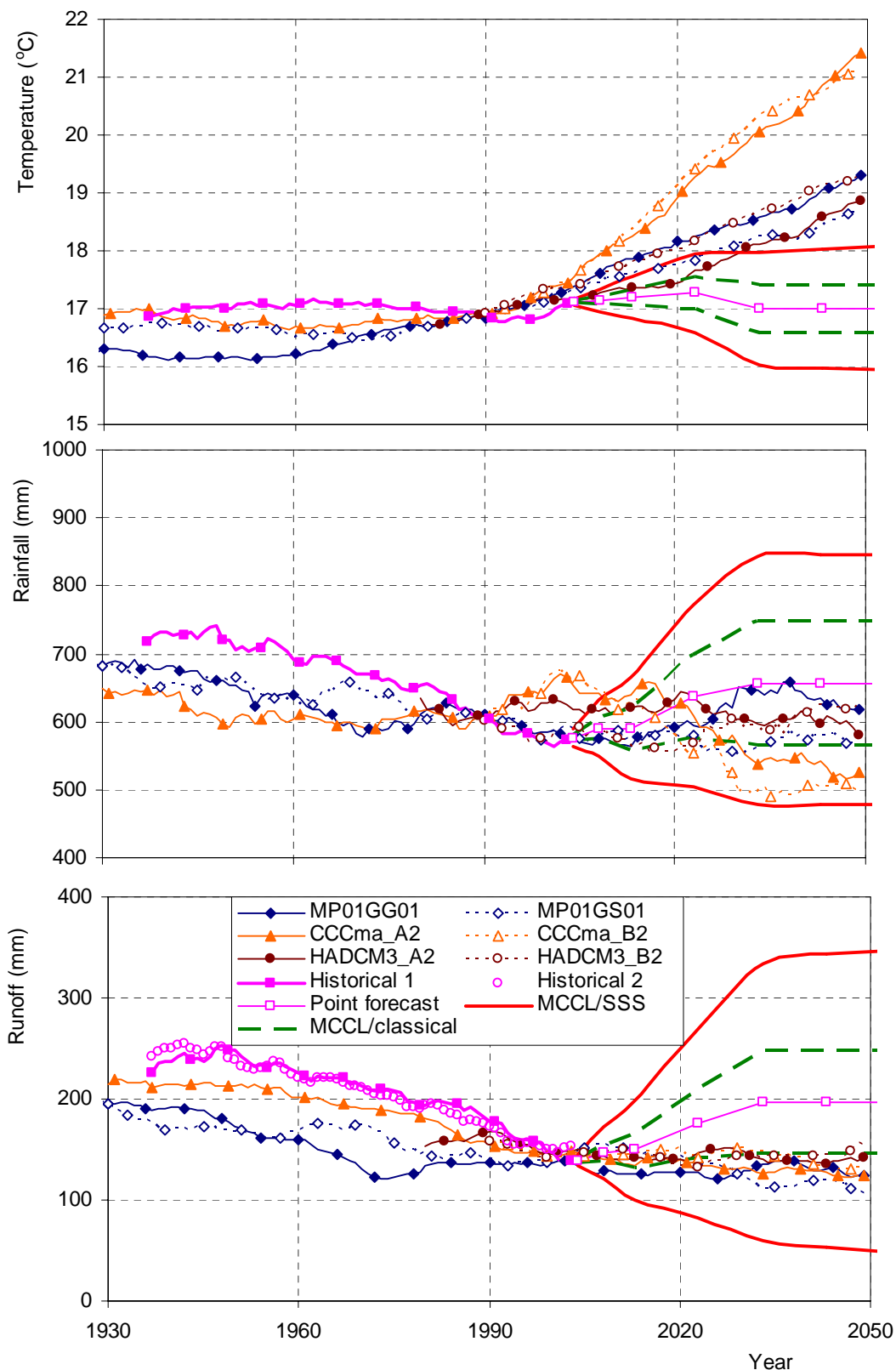


Figure 12 GCM predictions of past and future evolution of hydroclimatic processes at the Boeotikos Kephisos river basin vs. the uncertainty limits; “Historical 1” is measured data whereas “Historical 2” in the lower panel was produced by the hydrological model with historical rainfall and temperature inputs.

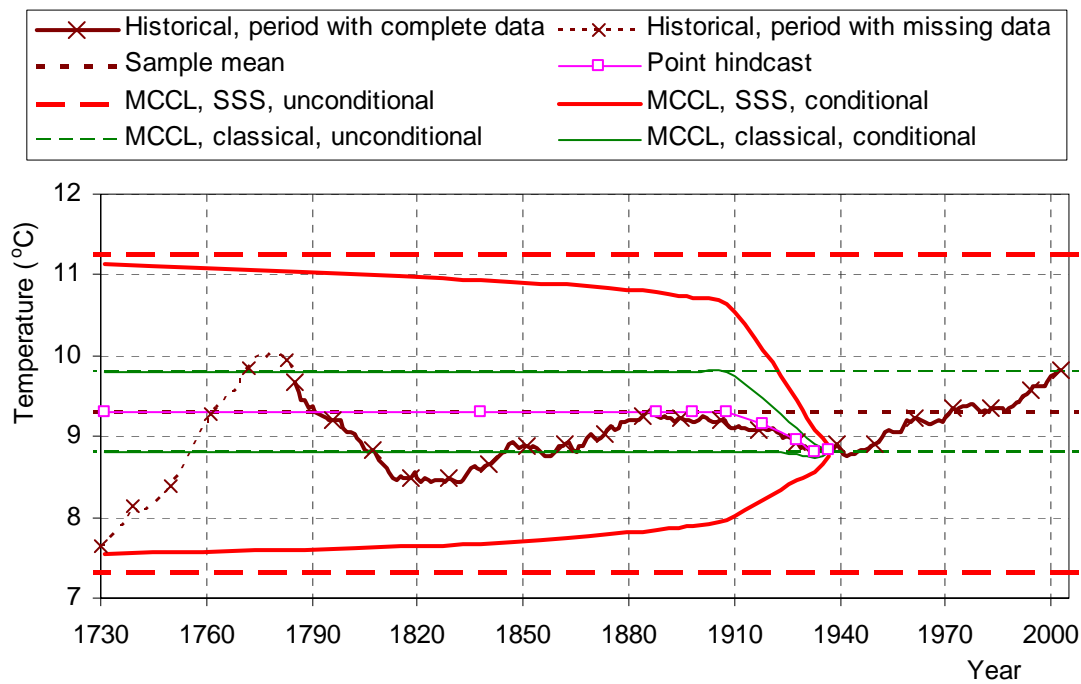


Figure 13 Historical climatic temperature at Berlin/Tempelhof and stochastic model hindcasts before 1938, based on the annual data of the period 1908-2003. The hindcasts are in terms of conditional point estimates and confidence limits (for $\alpha = \alpha' = 95\%$ and climatic time scale of 30 years).

Table 1. Observed sample statistics of the three long time series of the case study on an annual basis.

Sample statistic	Temperature (°C)	Rainfall (mm)	Runoff (mm)
Size, n	96	96	96
Mean, $x_0^{(n)}$	17.0	658.4	197.6
Standard deviation, s	0.72	158.9	87.6
Variation, $C_v = s/m$	0.04	0.24	0.44
Skewness, C_s	0.34	0.44	0.36
Lag-1 autocorrelation, r_1	0.31	0.10	0.34
Hurst coefficient, H	0.72	0.64	0.79

Table 2. IPCC scenarios for future climatic projections used in the study (from Leggett et al., 1992; Nakicenovic and Swart, 1999; Carter et al., 1999).

Label	Name	Description	Population in 2100 (billion)	CO ₂ concen- tration in 2100 (cm ³ /m ³)
S1	SRES A2	High population growth; high CO ₂ emissions	15.1	834
S2	SRES B2	Lower population; energy system predominantly hydrocarbon-based but with reduction in carbon intensity	10.4	601
S3	IS92a	In between S1 and S2	11.3	708

Table 3 Main characteristics of the GCM coupled atmosphere-ocean global models (http://cera-www.dkrz.de/IPCC_DDC/IS92a/Max-Planck-Institut/echam4opyc3.html; Flato and Boer, 2001; Gordon et al., 2000).

Label	Name	Developed by	Resolution (°) in latitude and longitude	Grid points, latitudes × longitudes
M1	ECHAM4/ OPYC3	Max-Planck-Institute for Meteorology & Deutsches Klimarechenzentrum, Hamburg, Germany	2.81 × 2.81	64 × 128
M2	CGCM2	Canadian Centre for Climate Modeling and Analysis	3.75 × 3.75	48 × 96
M3	HADCM3	Hadley Centre for Climate Prediction and Research	2.5 × 3.75	73 × 96

Table 4 General description of GCM output time series used in the study.

Label	Name	Model	Model inputs
T1	MP01GG01	M1	historical for 1860-1989; from S3 beyond 1990
T2	MP01GS01	M1	as in T1 but also considering the sulphate concentration
T3	CCCma_A2	M2	historical for 1900-1989; from S1 beyond 1990
T4	CCCma_B2	M2	historical for 1900-1989; from S2 beyond 1990
T5	HADCM3_A2	M3	historical for 1950-1989; from S1 beyond 1990
T6	HADCM3_B2	M3	historical for 1950-1989; from S2 beyond 1990

Table 5 Performance of climatic models before and after transformation

	Period of compari- son *	Bias (%) [†]	Overdisper- sion (%) [‡]	Coefficient of determ- ination	Coefficient of efficiency	Hurst coefficient
<i>Monthly temperature</i>						
MP01GG01	1908-89	10.5 / -2.7	-19.7 / -2.7	0.90 / 0.91	0.72 / 0.90	–
MP01GS01	1908-89	9.7 / -1.1	-20.6 / -0.9	0.90 / 0.92	0.72 / 0.91	–
CCCma_A2 & B2	1908-89	3.8 / -0.5	-45.9 / 1.0	0.58 / 0.92	-0.63 / 0.92	–
HADCM3_A2 & B2	1950-89	5.9 / -0.3	-40.7 / -1.8	0.76 / 0.92	0.04 / 0.91	–
<i>Monthly precipitation</i>						
MP01GG01	1908-89	-36.5 / -5.9	-53.2 / -7.4	0.06 / 0.17	-4.1 / -0.3	–
MP01GS01	1908-89	-36.7 / -5.7	-49.8 / -5.7	0.06 / 0.17	-3.6 / -0.3	–
CCCma_A2 & B2	1908-89	-15.7 / -7.3	-34.4 / -1.0	0.15 / 0.17	-1.2 / -0.2	–
HADCM3_A2 & B2	1950-89	-26.2 / 0.3	-22.2 / 0.2	0.16 / 0.12	-0.7 / -0.3	–
<i>Annual temperature</i>						
MP01GG01	1908-89	10.4 / -2.7	-30.1 / -3.2	0.00 / 0.00	-19.3 / -1.5	0.76 / 0.80
MP01GS01	1908-89	9.6 / -1.1	-27.1 / -1.1	0.00 / 0.00	-15.4 / -1.2	0.72 / 0.71
CCCma_A2 & B2	1908-89	3.5 / -0.5	-59.5 / 23.0	0.00 / 0.00	-12.1 / -0.7	0.51 / 0.49
HADCM3_A2 & B2	1950-89	5.6 / -0.5	-13.3 / 28.9	0.06 / 0.06	-5.9 / -1.0	0.86 / 0.86
<i>Annual precipitation</i>						
MP01GG01	1908-89	-36.5 / -5.9	-46.7 / -8.6	0.00 / 0.00	-12.5 / -1.3	0.47 / 0.50
MP01GS01	1908-89	-36.6 / -5.8	-34.6 / -2.3	0.01 / 0.01	-8.2 / -0.9	0.52 / 0.46
CCCma_A2 & B2	1908-89	-15.9 / -7.6	-36.9 / 0.8	0.01 / 0.00	-4.1 / -1.1	0.45 / 0.45
HADCM3_A2 & B2	1950-89	-26.0 / 0.4	-25.3 / 4.8	0.01 / 0.00	-4.5 / -0.8	0.42 / 0.49

* With historical time series

[†] Ratio of difference of means to historical mean[‡] Ratio of difference of standard deviations to historical standard deviation

Table 6 Performance of runoff estimated by the hydrological model using temperature and rainfall data, either historical or from transformed GCM outputs; annual scale.

	Period of comparison *	Bias (%) [†]	Overdispersion (%) [‡]	Coefficient of determination	Coefficient of efficiency	Hurst coefficient
Historical [§]	1908-2003	2.4	-9.9	0.68	0.59	0.81
MP01GG01	1908-2003	-22.7	-38.9	0.02	-3.0	0.91
MP01GS01	1908-2003	-17.3	-42.4	0.01	-3.3	0.89
CCCma_A2 & B2	1908-2003	-7.3	-38.3	0.12	-1.6	0.86
HADCM3_A2 & B2	1950-2003	-23.6	-43.1	0.01	-13.4	0.80

* With historical time series

[†] Ratio of difference of means to historical mean

[‡] Ratio of difference of standard deviations to historical standard deviation

[§] Hydrologic model output with historical temperature and rainfall