Nonstationarity versus scaling in hydrology

Demetris Koutsoyiannis

Department of Water Resources, Faculty of Civil Engineering, National Technical University of Athens, Heroon Polytechneiou 5, GR-157 80 Zographou, Greece (dk@itia.ntua.gr)

Abstract The perception of a changing climate, which impacts also hydrological processes, is now generally admitted. However, the way of handling the changing nature of climate in hydrologic practice and especially in hydrological statistics has not become clear so far. The most common modelling approach is to assume that long-term trends, which have been found to be omnipresent in long hydrological time series, are "deterministic" components of the time series and the processes represented by the time series are nonstationary. In this paper, it is maintained that this approach is contradictory in its rationale and even in the terminology it uses. As a result, it may imply misleading perception of phenomena and estimate of uncertainty. Besides, it is maintained that a stochastic approach hypothesizing stationarity and simultaneously admitting a scaling behaviour reproduces climatic trends (considering them as large-scale fluctuations) in a manner that is logically consistent, easy to apply and free of paradoxical results about uncertainty.

Keywords Climatic variability; Climatic changes; Hurst phenomenon; Hydrological design; Hydrological statistics; Nonstationarity; Risk; Scaling; Statistical estimates; Statistical tests; Uncertainty.

1. Introduction

Όλα τριγύρω αλλάζουνε κι όλα τα ίδια μένουν ...

Μανόλης Ρασούλης

(Everything around changes, yet everything remains the same ...

""""""""""""""Manolis Rasoulis)

As the development of hydrological science and technology relies greatly on measurements of the natural processes, its historical evolution can be paralleled to the hypothetical example of an experimentalist investigating in a laboratory a law relating two quantities x and y, as depicted in Figure 1 (adapted from Xanthopoulos, 1974). In a first phase, the observation window is too small (Window A in Figure 1) and the experimentalist may conclude that the quantity y is not related to x and it may have a constant value a plus/minus an error term. In a second phase, when the observation window becomes wider (Window B in Figure 1), a clear relationship between x and y emerges, which may be modelled by a parabolic law. In a third phase, the window becomes even wider (Window C), and the parabolic law is abandoned in favour of a cosine law. Probably, in a next phase even the cosine law may prove inadequate and a more complex law could be formulated.
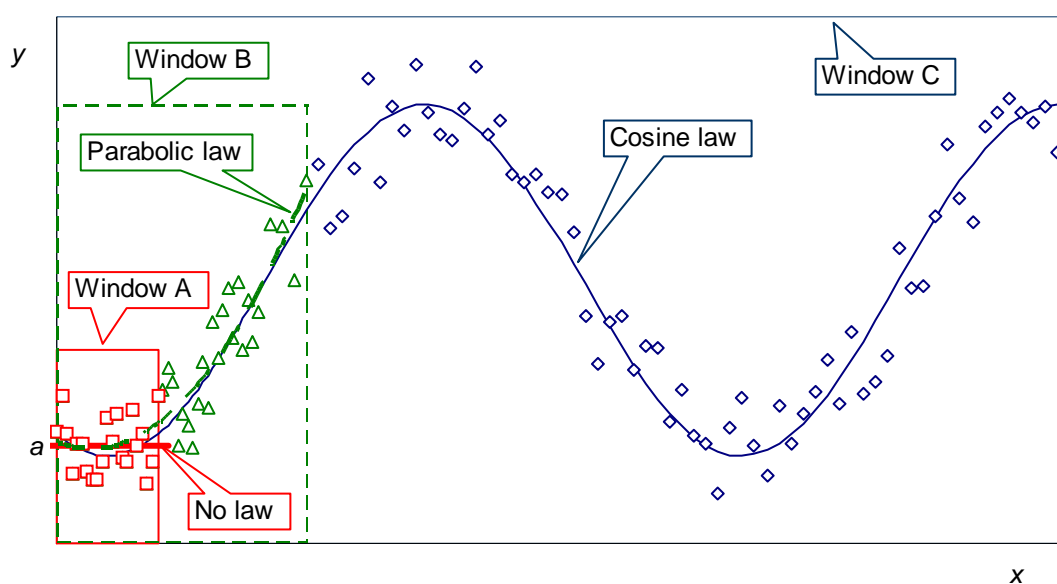

Figure 1. Depiction of a hypothetical example of successive experimental attempts to discover a law relating quantities x and y.

The job of a hydrologist is much more difficult than that of a laboratory experimentalist. For, the hydrologist cannot repeat any experiment or expedite the rate of experiments to quickly investigate a broader range of variations; instead, the hydrologist should keep observing until nature will produce a wider range of phenomena. The evolution of a hydrological process is unique.

Having moved from the laboratory environment of the experimentalist to the natural environment of the hydrologist, let us assume that the quantity x is time, expressed in discrete units of years, and the quantity y is, for instance, the annual runoff of a certain river basin. Obviously, the longer is the observation period, the more truthful is the picture of the runoff variation through time. It is natural to expect that, as time goes by and the observed time series becomes longer, our perception of the runoff process will change and become more and more accurate.

In the 1950s, only limited hydrological time series were available worldwide and these were short, with typical length 10-50 years. A typical situation is shown in Figure 2(a), where the time series of the annual runoff of the Boeoticos Kephisos River basin of the period 1907-1950 is plotted. The location of the river is to the north of Athens (Greece) and its basin is about 2 000 km^2. The annual runoff has been expressed in equivalent water depth in the catchment. The natural interpretation of this graph is to hypothesize a stable multiyear behaviour with annual random fluctuations around a constant mean (plotted as a thick horizontal line in Figure 2(a)).

Fortunately, and despite of the present time tendency to abandon hydrometric stations to reduce costs, the discharge measurements on this river continue until today; firstly owing to the convenient access and simple trapezoidal geometry of the river section, and secondly owing to the importance of the river which is part of the Athens water supply system. So, this runoff time series, which is depicted in complete length in Figure 2(b), has become the longest runoff sample in Greece. The whole picture now reveals behaviour quite different from the one of Figure 2(a). The mean seems to be not stable, but there is a long lasting falling trend extended from 1921 to present time (thick line AB in Figure 2(b)). What caused this trend? There is a simple answer, illustrated in Figure 2(c): A similar trend in rainfall. We

have performed an experiment (Koutsoyiannis and Efstratiadis, 2004) using a detailed hydrological model of this river basin (Rozos et al., 2004) fed by the historical rainfall. The model was able to reproduce perfectly the falling trend in runoff, and this is a sufficient reason to say that we have an explanation for that falling trend.
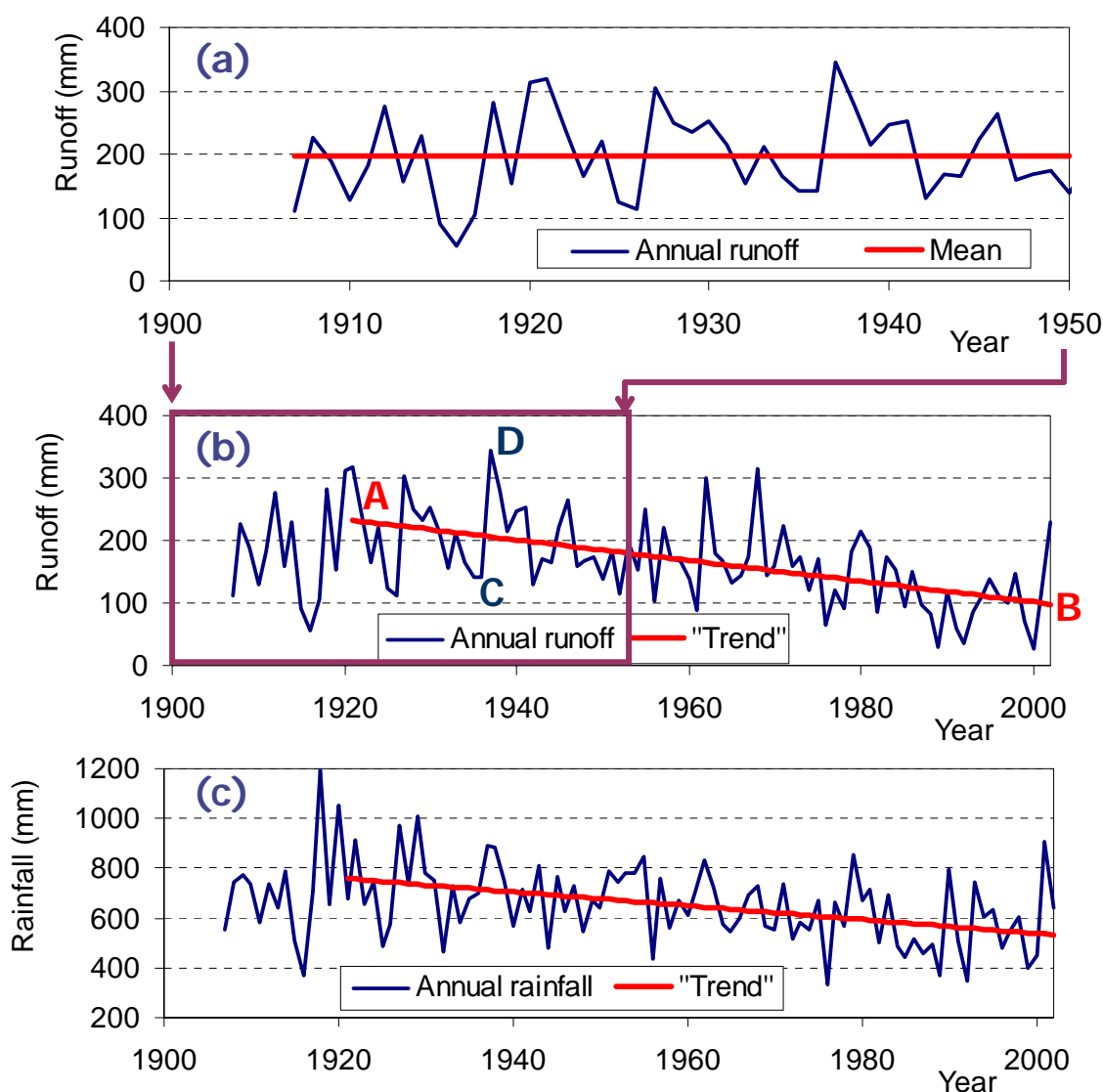

Figure 2. Plots of (a) the time series of annual runoff of the Boeoticos Kephisos River basin, Greece, between the years 1907-1950; (b) the same time series extended to year 2002; (c) the time series of rainfall at the same basin for the period 1907-2002.

Naturally, the next question is: What caused the trend in rainfall? This is much more difficult to answer. To study this question we must move from the hydrological basin to the atmosphere. At that time we must also leave the catchment scale and consider the global

(planet) scale, as the atmosphere has no borders like the divide of a catchment. For a hydrologist this may be too difficult or even impossible, and thus the task should be addressed to other scientists, such as climatologists. In climatology, it is well known that such a change in a climatic variable is not a unique behaviour at this very location, i.e. it is not an exception but it is the rule. This idea was endorsed by more recent evidence provided by paleoclimatological studies.

The hydrologist, however, has other important questions, whose answers cannot wait until climate modellers are able to explain the specific behaviour of rainfall at this catchment or at the global scale. These questions deal mainly, not with the explanation of the past, but with prediction of the future for the purposes of the design and management of hydrosystems. Undoubtedly, a concrete deterministic prediction of future evolution of the process of interest must be regarded as an impossible task. Therefore, a prediction of future uncertainty is the best target that a hydrologist can hope to accomplish. However, such a prediction depends seriously on the information of the past. Thus, we can anticipate that the estimate of future uncertainty will be much more faithful in case that the complete time series of Figure 2(b) is available than in case of the short information of Figure 2(a). Most important, the modelling approaches that must be followed in each of the two cases will be different. The stable mean in Figure 2(a) implies that the uncertainty estimation can be approached by the classical hydrological statistics (i.e. estimating standard deviation, fitting a certain distribution, estimating confidence limits, etc.). But, what about the changing mean case of Figure 2(b)? Obviously, the interpretation and modelling of the changing mean affects seriously the design and management of hydrosystems. Here two different modelling approaches have been followed: (1) the nonstationarity approach, and (2) the scaling approach. These are the theme of this paper and are discussed in detail in the next two sections. As will be demonstrated, the two modelling approaches differ dramatically in their logic and interpretation of natural phenomena, and most importantly, in the estimates of uncertainty they produce, which according to the second approach are much higher than in the first. Thus, the comparison of the two modelling approaches, which is the target of the paper, is not merely a theoretical issue but it has significant practical consequences, given that, when dealing with the uncertain

natural behaviour, an engineer's duty is to quantify uncertainty as faithfully as possible and to incorporate this quantification in management decisions.

2. The nonstationarity approach

2.1 Typical procedure

This is the most common modelling approach followed in hydrologic practice when a pattern like that in Figure 2(b) is brought to light. In fact, this approach is so common that most hydrological books mention it as the only modelling approach to follow in such a case. The rationale of the typical approach has been expressed by Eagleson (1970, p. 155-156), who considered each hydrologic process to be composed of three parts:

> *"1. A deterministic part which is periodic and results from natural physical periodicities.*
>
> *2. A deterministic part which is aperiodic. This is commonly called "trend" and may be considered to be due to a gradual temporal change in the physical parameters of the processes controlling weather.*
>
> *3. A stationary random part."*

Here we can observe several tacit assumptions that we will discuss further in subsequent sections. Among these are the identity "random = stationary", the dichotomies "hydrologic process = deterministic + random", "deterministic = periodic + aperiodic" and "aperiodic = deterministic + random" (as it is well known that random is aperiodic too). This description of hydrological processes appeared earlier (Chow, 1964, p. 8.10), albeit in less definite context, and later (e.g. Yevjevich, 1972, pp. 17-20; Kottegoda, 1980, p. 26), sometimes with few changes. For example Salas et al. (1980, p. 33) added a fourth component related to "almost periodic" changes.

Technically, the modelling approach followed in hydrologic practice includes the following steps whose details can be found in most widespread hydrological texts (e.g. Dawdy and Matalas, 1964; Kottegoda, 1980, pp. 22-34; Salas, 1993, pp. 19.5-19.7, 19.17-19.19) and partly in business statistics texts (e.g. Freund et al., 1988, p. 583-598):

1. Perform statistical tests to diagnose whether the trend is statistically significant or not.

2. If the tests confirm that the trend is significant, continue with the following steps, otherwise treat the series as being random.

3. Fit a function to the trend (such as a linear expression of time) and perform additional statistical tests to verify that the parameters of the function are statistically significant.

4. Call the trend a deterministic component of the time series and call the time series a nonstationary one.

5. "Detrend" the time series (i.e., subtract the trend from the original series) and use the "detrended" series (the residuals) for assessment of uncertainty or for modelling of the process.

It is the opinion of this author that this modelling approach is contradictory in its rationale and its terminology, in its separate steps and as a whole. As a result, it may imply misleading perception of the phenomena and uncertainty estimation. This opinion is explained in next subsections, separately for each of the steps of the approach.

2.2 Use of statistical tests

Typically, a trend is detected using the nonparametric Kendall's test (e.g. Kottegoda, 1980, p. 32) with test statistic $\tau := 4p/[n(n-1)] - 1$, where p is the number of pairs of observations (x_j, x_i; $j > i$) in which $x_j < x_i$. In a random series, τ has mean 0, variance $2(2n + 5)/[9n(n-1)]$, and distribution converging rapidly to normal. In the case of Boeticos Kephisos runoff, application of the test results in $\tau = 0.40$ with standard deviation 0.077, and eventually Kendall's test results in rejection of the null hypothesis that a trend does not exist for an attained significance level as low as 8.8×10^{-8} (Koutsoyiannis, 2003); obviously, this is regarded as enormously high statistical evidence that a trend really exists.

However, there are two problems here, which make the result incorrect. Firstly, it is known that the validity of confirmatory tests is based on the assumption that the investigator developed the hypothesis prior to examining the data (Hirsch *et al.*, 1993, p. 17.5). In this case, the hypothesis was formed after examining the data and seeing that there is a downward

trend. This problem along with other misuses of statistics in climate research has been discussed by von Storch (1995). Secondly, the null hypothesis contains the sub-hypothesis that the series is random. Specifically, the variance of the test statistic τ is calculated theoretically assuming that the sample examined is purely random. Obviously, the random prototype is not an appropriate model for a hydrological series. It is well known that hydrological series incorporate some stochastic structure, expressed by nonzero serial correlation. The neglecting of the effect of serial correlation to trend detection is put by von Storch (1995) as another standard example of misuse of statistics. Given these two fundamental errors, the rejection of such a null hypothesis does not prove the existence of a trend. This would be clarified further in subsection 3.3.

2.3 The fitting of a function

Typically, to express mathematically a trend, a polynomial function of time is used (e.g. Dawdy and Matalas, 1964, p. 8.81; Kottegoda, 1980, p. 26; Salas et al., 1980, p. 45). The first order polynomial, i.e. a linear function, is the simplest possible and therefore the most widely used to model hydrological trends. It also enables easy statistical testing of the significance of the estimated slope of the line, so it provides additional means of testing the existence of a trend (e.g. Kottegoda, 1980, p. 32). In our example, the significance level attained by such a test is of the same order of magnitude as in the Kendall's test described above, so the test concurs with the existence of a statistically significant trend.

Apparently, the problems already mentioned in subsection 2.2 about statistical testing apply also here. However there are additional, more fundamental problems in this case. The important question here is whether any fitted equation has physical meaning or not. Without doubt, in the laboratory experiment discussed in the Introduction (which may have indirectly served as a prototype in identifying trends in hydrology), the fitting of an equation (e.g. a quadratic law in the case of Window B or a cosine law in case of Window C) must have a physical meaning. But, in the case of the Boeoticos Kephisos runoff the situation is dramatically different. To be more explicit, in the laboratory case the experiments may be repeatable, and most probably the points plotted in Figure 1 originate from numerous

experiments rather than from a single one. This enhances our belief to the law. Physically, the law expresses the dependence of quantity y on x, whereas the observed deviations of the points y from the fitted law $y(x)$ may be regarded either as measurement errors or as the effects of factors other than quantity x. On the other hand, in the runoff case no repetition of measurement can be executed. Besides, the line AB in Figure 2(b) does not describe the dependence of runoff on time and the deviations of the rough line plot of the time series in Figure 2(b) cannot be regarded as measurement errors. Rather, the rough line itself expresses the dependence of runoff on time, but even the rough line cannot be called a physical law, even though, undoubtedly, it contains a certain physical meaning.

2.4 The notion of determinism

As mentioned in subsection 2.1, a simple function fitted to a data series, such as the linear function represented by line AB in Figure 2(b), is typically (according to the approach discussed herein) considered to be a deterministic component of the time series whereas the deviations of the time series from this line are supposed to be the random component. Here a simple question arises: Why consider the line AB in Figure 2(b) a deterministic law and the deviations from this line random? Why, for instance, line CD, which connects two consecutive points of the time series in Figure 2(b), is more random and less deterministic than line AB? After all, AB was established by statistical means, whereas at the same time, CD is a result of measurement of a physical process.

   If the line AB were obtained, for instance, as a result of a hypothetical physically based model, which would be able to produce a trend for the future based on data of the past, the characterization "deterministic" would be justified. But this is not the case at present (e.g. Koutsoyiannis and Efstratiadis, 2004) and probably it will be not the case in the foreseeable future. Thus, the statistical approach described above remains the main tool to identify and describe such patterns. But given the facts that: (1) they were identified only *a posteriori* by statistical processing of historical data, and (2) they are not verifiable by any experiment (due to the uniqueness of the natural process), these patterns do not deserve the characterization "deterministic".

Here it should be clarified that the proposed point of view does not deny determinism in the evolution of hydrological processes. Rather, it tries to show that the modelling of the trend line as a deterministic law is not consistent with the modelling of the deviation from this law as a random or stochastic component. In other words, it is maintained that, if the modelling of line CD in Figure 2(b), or that of the complete rough line plot depicting the time series, is to be done in stochastic terms, then the modelling of the trend line AB should also be done in stochastic terms. Here it should be emphasized that the modelling of a physical system using a stochastic model is not a matter of characterizing its nature or structure: after all, every macroscopic physical system can be regarded as deterministic in its structure (here we must exclude microscopic quantum systems, in which indeterminism may be intrinsic). But there are cases where determinism does not help to study and predict complex macroscopic systems and in these cases it is better to use stochastic, probability-based, models. To quote von Plato (1994, p.15):

> "*In classical physics probabilities are basically nonphysical, epistemic additions to the physical structure, a 'luxury' as von Neumann says, while quantum physics, in contrast, has probabilities which stem from the chancy nature of the microscopic world itself. Epistemic probability is a matter of 'degree of ignorance' or of opinion, if you permit*".

2.5 The notion of nonstationarity

As mentioned in subsection 2.1, according to the modelling approach examined, a time series, in which a trend is identified, is regarded as nonstationary. Since similar trends have been identified in many long time series representing hydrologic processes, there is a tendency to characterize hydrologic processes as nonstationary. But is this justified?

Before we study this question, we should clarify the notions and terms of a "process" and a "time series" (see explanatory sketch in Figure 3). Firstly, it is useful to distinguish the meaning of the term "process" in the mathematical theory of stochastic processes and in the context of real world (physical, chemical, biological, social, etc.), processes. In Greek, to avoid confusion because of the different meanings, different terms are used in the two cases.

Thus, "stochastic anhelixis" (στοχαστική ανέλιξις) is a "stochastic process" and "physical diergasia" (φυσική διεργασία) is a "physical process". (The term "anhelixis" is etymologized from "helix" (έλιξ), here related to evolution, whereas the term "diergasia" is etymologized from "ergon" (έργον), i.e. work, as in "en-erg-y", "ergo-nomical", or "erg-odicity".)

Secondly, we must distinguish a time series $x(t)$ from a process $X(t)$. The former is a specific trajectory or path of the latter. Thus, in a real world process (diergasia), a time series is a series of observations of the process. In a stochastic process (anhelixis), a time series is a specific realization of the process. Consequently, in both cases a time series is a sequence of numbers (values). In contrast, a stochastic process is a family of random variables (typically, indexed by time; e.g. Taylor and Karlin, 1984, p. 4; Weisstein, 1999-2005b). Here, it should be mentioned that this terminology and the distinction of a process and a time series is prevailing in most texts, either from the disciplines of stochastics (e.g. Box et al., 1994, p. 7), dynamical systems (e.g. Brock and Potter, 1992; Theiler et al., 1994), and hydrology (e.g. Chow, 1964, p. 8.9) but it is not universally accepted, which creates confusion. For example, in Kotz and Johnson (1988) and Weisstein (1999-2005a) the term time series is used as essentially synonymous to stochastic process. Some hydrological texts may make a distinction of the two terms but on other grounds (for example, Clarke, 1998, distinguishes two classes of models, time series models and stochastic models – however, according to the above prevailing definition both classes would be called stochastic processes). Some books in probability and stochastics (e.g. Karlin and Taylor, 1975; Taylor and Karlin, 1984 Papoulis, 1991) avoid the use of time series, probably to prevent confusion, favouring the term "sample function" instead. Here we use both terms, as in Figure 3, with the prevailing definitions stated above.
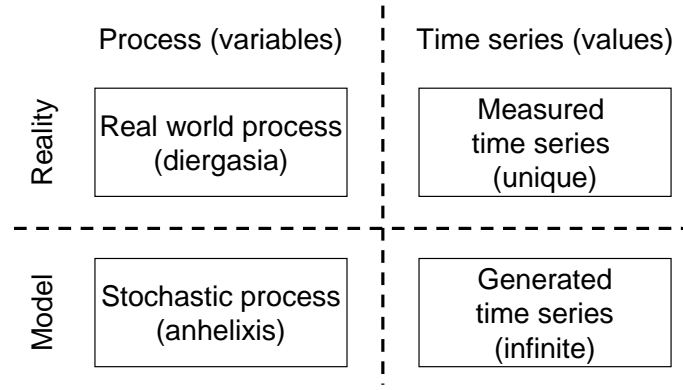
	Process (variables)	Time series (values)
Reality	Real world process (diergasia)	Measured time series (unique)
Model	Stochastic process (anhelixis)	Generated time series (infinite)

Figure 3. Schematic for the clarification of the terms "process" and "time series".

A stochastic process implies infinite realizations. Thus, for example, in a stochastic process (anhelixis) $X(t)$ we can think of the mean of the process, defined as

$$\mu(t) = \int_{-\infty}^{\infty} X(t)\, dF_{X(t)}(x) \tag{1}$$

where $F_{X(t)}(x)$ is the probability distribution function of $X(t)$ at time t. Indeed, behind this definition lies the assumption of infinite potential realizations of $X(t)$, therefore we call $\mu(t)$ the ensemble mean. Then we can define *stationarity* of the mean assuming that $\mu(t)$ does not vary with t [i.e., $\mu(t) = \mu$], etc.

Now let us come to the notion of stationarity in physics. Here, the term stationarity may have different meanings depending on the context. Thus, in thermodynamics we have the notion of a stationary system, in which the changes of kinetic and potential energies are zero (Çengel and Boles, 2002, p. 168). This notion is fundamentally different from, and not related to, the notion of stationarity discussed here. In dynamical systems we have the notion of a *stationary point*: we call a point x^* stationary if it is invariant under the dynamical evolution, i.e. $\mathbf{S}_t(x^*) = x^*$, where $\mathbf{S}_t$ denotes the dynamical transformation of the system (Lasota and Mackey, 1994, p. 191). In this case we would call stationary a process that is time invariant: $X(t) = x^*$, which however is a trivial non-interesting case. Also in dynamical systems, we have the more general notion of stationary densities (Lasota and Mackey, 1994, pp. 41, 417) whose definition is practically equivalent to that of stationary stochastic processes. In this case, a density is defined in terms of a measure space, which is a generalization of a probability

space. Again, as in stochastic processes, infinite potential realizations of the process are implied (thus, we can speak of a density of potential outcomes), even though the system dynamics may be totally deterministic (a typical example of such dynamics, also used by Lasota and Mackey, 1994, for an intuitive presentation of the notion of stationary density, is the logistic map).

In conclusion, the notion of stationarity is closely associated with a model (rather than a real world system) either deterministic or stochastic, which implies an ensemble of infinite realizations. Now, let us come to a real world system *per se* (not a model of the system) such as the Boeoticos Kephisos River or its catchment. Obviously, this system has a unique evolution $X(t)$ in time t. We can measure the state of the system at certain times and obtain a time series of measurements $x(t)$, which is unique. Can we define the ensemble mean? The answer must be, no. What we can define is the notion of temporal mean, i.e.,

$$m(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x(t)\, dt \qquad (2)$$

which is fundamentally different from the ensemble mean $\mu(t)$. There is no way to assign a unique quantity, say $m(t)$, at a certain time instant t, unless we identify $m(t)$ with $x(t)$. So, can we define stationarity and nonstationarity for the real world process? The author's opinion is, no. That is to say, the notion of stationarity (and nonstationarity) is well defined within the theory of stochastic processes (being called a stochastic process stationary if its statistical properties are invariant to a shift of time and nonstationary if they are *deterministic* functions of time; Papoulis, 1991, p. 297), or even in mathematical models in general, but may not have any meaning outside of mathematics. In the case that we would identify $m(t)$ with $x(t)$ and postulate stationarity, we would call stationary a process that is time invariant: $x(t) = x^*$, which is a trivial case as discussed above.

But what about a given time series? Can it be characterized stationary or nonstationary? The typical answer in hydrological texts is positive. For example, Salas (1993, p. 5) states that "A hydrologic time series is stationary if it is free of trends, shifts or periodicity". In other disciplines we meet more careful statements, e.g. in Theiler et al. (1994, p. 453): "While

stationarity has a clear-cut meaning for a stochastic process, it is a fuzzier concept when applied to a time series". Probably, however, fuzziness could be removed by distinguishing time series generated by a model (anhelixis) or a real word process (diergasia). In the first case, if we know that the given time series has been generated by a certain stochastic process, which is known to be stationary or nonstationary, we could convey this property of the process to the time series and call it respectively stationary or nonstationary. But in the second case, maybe there does not exist any objective way to characterize it stationary or nonstationary. Even in the first case, in a time series generated from a stochastic process, we cannot say that the time series is stationary or nonstationary unless we know the generating process. In fact, any short time series (theoretically, any time series with finite length) can be generated by infinite stochastic processes, stationary and nonstationary.

In conclusion, a natural process (in our case a hydrological process) and a time series generated by measuring this process cannot be characterized as stationary or nonstationary because such a classification does not have a meaning. A stochastic process is always a model of a physical system; it is not a physical process *per se*. Stochastic processes allow convenient description and handling of several physical systems. Also, they allow the use of probabilistic description of uncertainty and the notion of infinite realizations, in place of the unique evolution of natural systems. A stochastic process that is used to model the physical process can be stationary or nonstationary. In most cases, this is primarily a matter of modelling convenience rather than a matter of faithful representation of the natural process. In other cases, nonstationarity of the process may be implied theoretically from its definition. This is the case for example in cumulative processes (such as a process modelling cumulative rainfall, where $X(t + \tau) \geq X(t)$ for any $\tau \geq 0$). Such a process is nonstationary by definition but it may have stationary intervals.

The opinion that a stationary model can be used to describe a hydrological process that is obviously affected by numerous time varying factors may seem strange to many. Therefore, we will discuss it further with the help of an example from rainfall modelling. Rainfall is affected by several mechanisms and events including the strangest ones such as volcano eruptions. Even if we ignore these complex influences and we only consider the most

common ones, like fronts, depressions and convection, it suffices to understand that the formation of a certain front is a deterministic process with a unique temporal existence; i.e. it happens today and not tomorrow or in a month. What will happen in a month is the formation of a different front, just as the Pinatubo eruption is different from the Santorini eruption. Does this imply nonstationarity? As explained before, we cannot characterize the physical process as stationary or nonstationary. However, we can model the rainfall process either as a stationary, cyclostationary or nonstationary stochastic process, depending on the modelling target, the additional information or tools we have available, and the modelling time scale (the relation of the time scale has been discussed by Klemeš, 2002). If we had a deterministic model that could predict the generation of fronts and depressions (or the volcano eruptions and their effects to rainfall), then we would model it as a nonstationary process, for example with varying mean $\mu(t)$ in time t, where the function $\mu(t)$ would be derived by the deterministic model. If we did not have such a deterministic model, we could assume that the occurrences of fronts, depressions, eruptions, etc. can be modelled, for instance, as random points in time. In such a case we would superimpose a stochastic model for storm occurrence and another model for each individual storm's structure, and it is very probable that the resulting compound stochastic process would be a stationary stochastic process.

This is the case for example in the Neymann-Scott and the Bartlett-Lewis point processes that have been used in stochastic rainfall modelling. If in each of these models the storm origins could be fixed by, say, a deterministic storm occurrence model, then the models would be nonstationary. If the storm origins are not known but assumed to obey a Poisson process, then the models are stationary. Thus, the stationarity is again a matter of 'degree of ignorance', not a matter of any essence in the physical process.

Obviously, solid information or knowledge (as opposite to ignorance) of a certain event affecting a process may justify the use of a nonstationary model. For example, to model river flow downstream of a dam we would use a nonstationary model with a shift in the statistical characteristics before and after the construction of the dam. Even in this case, however, caution is needed to understand and describe the causal physical mechanisms that affect the process; merely the occurrence of an event and the calculation of statistics may not suffice to

justify the use of a nonstationary model. A good counterexample, described by Klemeš, (2002), refers to the drop of the river Nile flows in the beginning of the 20th century, which was incorrectly attributed to the first filling of the old Aswan dam (1902-1903) whereas this reflected a similar drop in precipitation around the turn of the century over a vast region including much of the tropical Africa and the West Indies – a drop that can hardly be attributed to the Aswan dam.

2.6 The detrending of the time series

The above discussion about the use of the term "nonstationarity" for a time series that seems to have a trend, and the term "deterministic" (component) for this trend, is not merely an issue of terminology, mathematical logic, or philosophy. It has also significant practical consequences. More explicitly, if one follows the nonstationarity approach, i.e. calls the time series nonstationary and the trend deterministic, the next natural step is to separate randomness from deterministic behaviour, i.e., to detrend the series. In the example of Figure 2, during the period 1921-2002 the standard deviation, which is a measure of uncertainty, is 68.2 mm. If we detrend the series, the resulting standard deviation becomes 56.3 mm. This naturally means that the "deterministic" component explains an appreciable percentage 32% $[= 1 - (56.3/68.2)^2]$ of the process variability.

So, the practical consequence, according to this modelling approach, is that trends decrease uncertainty. This would be correct if the identified trend was indeed a deterministic component, so that one could apply to the future the same law that was established from past data. But is that the case in our example? Will this linear trend continue in the future? No one knows (in fact it is very unlikely because, if it continues, soon the runoff will become negative!). And if no one knows, we do not have reasons to believe that the uncertainty has indeed decreased by fitting this "deterministic" equation to the original time series and calling the time series nonstationary.

In conclusion, the nonstationarity approach implies several logical inconsistencies and finally yields a paradoxical impression of decrease of uncertainty. The only merit of this modelling approach is that, by detrending the original series, we obtain a time series that can

be regarded as a random sample (because it is free of trends), so that we can use classical statistics to process it further.

3. The scaling approach

3.1 The notion of scaling

The scaling modelling approach is the antipode of the nonstationarity approach. Instead of trying to adapt the series so as to obey classical statistics, it adapts classical statistics so as to be consistent with the observed behaviour of natural time series. It models the trend as a (large-scale) stochastic component and assumes stationarity. In fact, it does not require the separation of the time series into two or more components, so it does not attempt to detrend the original series. It admits that the existence of trends is the normal behaviour of real world time series. And it shows that trends increase dramatically the uncertainty rather than decrease it.

Before we proceed to the description of the scaling behaviour, it may be useful to re-examine our example time series, that of the Boeoticos Kephisos runoff, and compare it to a longer series, so as to acquire an idea of the existence and behaviour of trends on a longer term. For this purpose, we have chosen the longest available hydrological time series, the Nilometer series, which gives an index of the annual minimum water level of the Nile river for the years 622 to 1284 A.D. (663 observations, recently published by Beran, 1994 and available online from http://lib.stat.cmu.edu/S/beran). In Figure 4(a) we have reproduced the Boeoticos Kephisos runoff time series from Figure 2(b). Just below, in panel (b) we have plotted a hundred-year part of the Nilometer time series, where it can be seen that a trend similar to that of Boeoticos Kephisos runoff is present in the Nilometer time series between years 700-780 A.D. In panel (c) we have plotted the complete Nilometer series as well as its 5- and 25-year averages. Here, it can bee seen that trends (mostly viewed from the 25-year scale) exist through the entire span of the time series and their behaviour is irregular. More specifically, the downward trend of the years 700-780 A.D. changes to an upward trend just after 780 A.D.
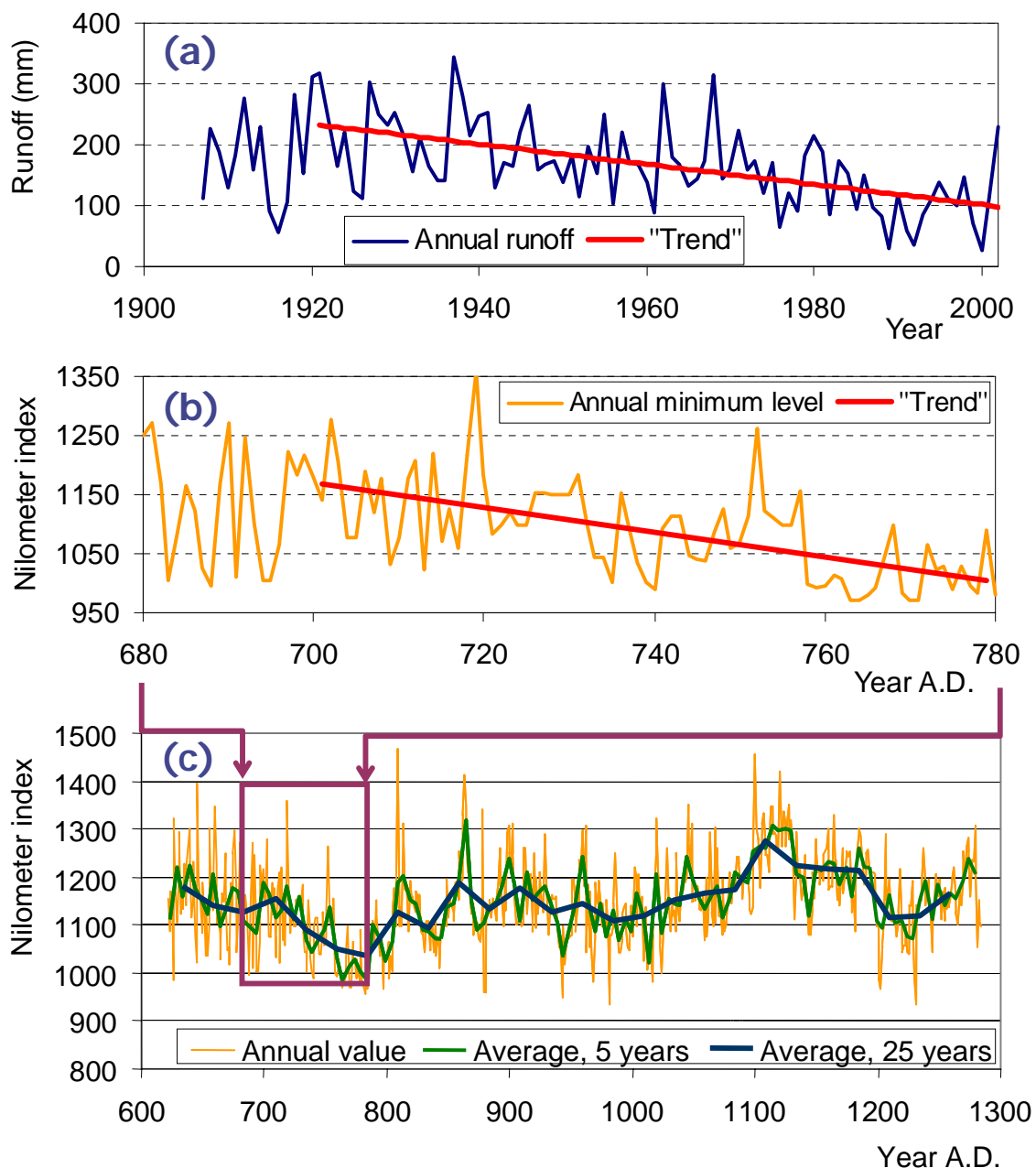
Figure 4. (a) Plots of (a) the time series of Boeoticos Kephisos annual runoff; (b) a hundred-year part of the Nilometer time series; (c) the complete Nilometer time series.

So, moving from the hundred-year observation window to 663-year window our perception of the process changes; this move can be paralleled to the shift from Window B to Window C in the initial example of Figure 1. Can we speak now about deterministic trends? Rather, instead of speaking of trends, it may be more consistent to speak of fluctuations at multiple time scales. And this harmonizes with the fact that climate changes on all time scales.

Now, in classical statistics we have the fundamental law

$$\text{StD}[\overline{X}_n] = \frac{\sigma}{\sqrt{n}} \qquad (3)$$

where $\overline{X}_n$ is the sample average estimated from a sample X_i of size n, i.e., $\overline{X}_n := (X_1 + \ldots + X_n) / n$, StD[ ] denotes the standard deviation of a random variable, and σ is the standard deviation of each of X_i (assuming stationarity, so that all X_i have common mean, standard deviation, etc.). Given a time series of sufficient length m, we can test in a simple way whether this law is fulfilled or not. To this aim, we can estimate from the time series the standard deviation StD[$\overline{X}_n$] for several values of n. Assuming $n = 1$, it is easily seen that StD[$\overline{X}_1$] ≡ StD[X_i], so this is the estimator of σ; for this estimate m data values are available. For $n = 2$ (and assuming for simplicity that the series length m is even) we can construct ($m/2$) samples $\overline{X}_2^{(1)} = (X_1 + X_2)/2$, $\overline{X}_2^{(2)} = (X_3 + X_4)/2$, …, $\overline{X}_2^{(m/2)} = (X_{m-1} + X_m)/2$. From these we can estimate StD[$\overline{X}_2$]. Proceeding this way, we can estimate StD[$\overline{X}_n$] for n up to, say, $m/10$ (in order to have 10 samples for the estimation of standard deviation). Plotting the estimate of standard deviation StD[$\overline{X}_n$] versus n (preferably on a logarithmic plot) we can test graphically the validity of the statistical law (3). Such plots for the Boeotikos Kephisos and the Nilometer time series are given in Figure 5. The non fulfilment of law (3) is more than clear. Instead, both plots follow a more general law of the form

$$\text{StD}[\overline{X}_n] = \frac{\sigma}{n^{1-H}} \qquad (4)$$

where H is a constant between 0.5 and 1. As shown in Figure 5, $H = 0.79$ for Boeoticos Kephisos and $H = 0.85$ for Nile.

Now, instead of modifying our time series to make it a random series (e.g. by detrending it) in order to fulfil the statistical law (3), it may be more wise to abandon the postulate of pure randomness and devise a stochastic process that has the property (4) from the outset. This is mathematically possible and such a process can be called a simple scaling stochastic process (SSS process). In fact, (4) can serve as a definition of an SSS process, sufficient for the purpose of this paper. (For more generalized definition of the process in discrete time in terms of probability distribution rather than merely standard deviation see Koutsoyiannis, 2002, 2003; for a definition in continuous time see Abry et al., 1995 and Beran, 1994).
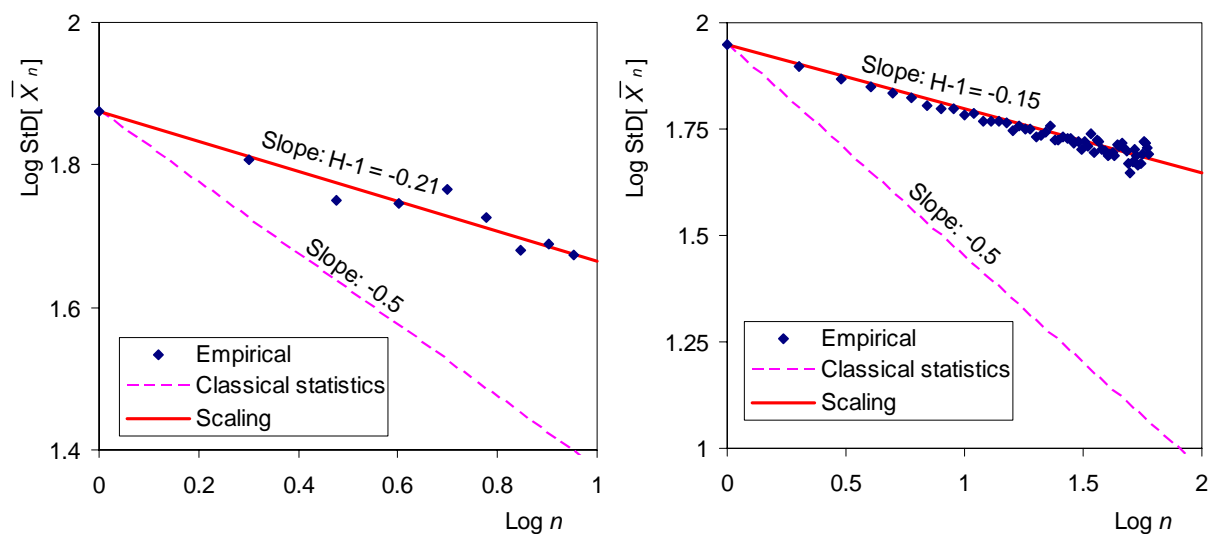
Figure 5. Logarithmic plots of standard deviation versus scale for the time series of (left) Boeoticos Kephisos annual runoff; (right) Nilometer.

The statistical behaviour expressed by equation (4) is none other than the Hurst phenomenon, known in hydrology since more than half a century ago, and the constant H is the Hurst exponent. Hurst (1951) who discovered this behaviour used a different formulation for the law, based on the so called rescaled range. The formulation in terms of standard deviation at the time scale n, as in equation (4), is much simpler yet equivalent to that of Hurst (see theoretical discussion by Beran, 1994, p. 83, and practical demonstration by Koutsoyiannis, 2002).

The Hurst or scaling behaviour has been found to be omnipresent in several long time series from hydrological, geophysical, technological and socio-economic processes. Thus, it seems that in real world processes this behaviour is the rule rather than the exception. The omnipresence can be explained based either on dynamical systems with changing parameters (Koutsoyiannis, 2005b) or on the principle of maximum entropy applied to stochastic processes at all time scales simultaneously (Koutsoyiannis, 2005a).

3.2 Typical procedure

In subsection 2.1 it was described that the typical procedure to handle a trend by the nonstationarity approach contains a handful of steps. In the case of the scaling modelling approach, the typical procedure is even simpler; it includes three steps:

1. Construct a double logarithmic plot of standard deviation versus scale.
2. Estimate the Hurst coefficient H from the slope of this plot.
3. If $H \approx 0.5$ treat the series using classical statistics; otherwise use modified statistics appropriate for SSS processes.

The first two steps have been already described by means of the example in Figure 5. If more accuracy in estimation of H is needed, improved algorithms (Beran, 1994; Koutsoyiannis, 2003) may be used in step 2, which may simultaneously provide a better estimate of σ, the standard deviation at the basic scale. The modified statistics (e.g. statistical estimators and tests) required for step 3 are described in Koutsoyiannis (2003). If there does not exist an analytical solution for the statistical properties of a certain statistic in step 3 (e.g. for Kendall's τ statistic discussed in subsection 2.2), stochastic simulation is the method of choice. This requires generation of ensembles of synthetic series from an SSS process, which can be done using simple generators (appropriate for spreadsheet calculations) as discussed in Koutsoyiannis (2002).

3.3 Relation of the scaling behaviour to trends in hydrologic series

There are two simple intuitive ways to verify the close relation between the scaling behaviour and trends in hydrologic series. The first is to detrend an observed time series and reconstruct a standard deviation versus scale plot such as that of Figure 5. It will be seen that the scaling behaviour disappears and the slope of the arrangement of empirical points approaches 0.5, as predicted by the classical statistical law (3). The second is to generate some synthetic series with a Hurst coefficient greater that 0.5 (say 0.8) and plot them. It will be seen that falling or rising trends are very common.

The second method can be extended in a more systematic manner to perform statistical tests based on stochastic simulation. Let us consider again Kendall's τ statistic described in subsection 2.2, by means of which it was decided that there exists a statistically significant trend in the Boeoticos Kephisos runoff series. The same test and the same statistic can be used in an SSS framework as well. The difference is that the standard deviation of τ is no longer given by the classical statistical formula (subsection 2.2). The easiest way to calculate the standard deviation of τ is to perform simulation. In this respect, in Koutsoyiannis (2002) an ensemble of 100 time series was generated, each with $n = 91$ and $H = 0.78$, and in each of these series that 78-year period which gave the maximum value of τ (in absolute value) was located (note that fewer years of data were available then, which explains the slight differences in the length of trend period and the value of H). The standard deviation of τ over the 100 series was 0.252 (more than three times greater than the value 0.077 that was calculated by the classical statistical formula) and the attained significance level of Kendall's test was 5.5% (almost six orders of magnitudes higher than the one of classical statistics!). This means that the trend is not statistically significant at the usual 5% significance level. In simpler words, trends are so common in time series generated by an SSS process that Kendall's test deems a 78-year falling trend in a 91-year data series a regular behaviour.

The relation of trends, or better of large-scale fluctuations, to the scaling behaviour, or equivalently to the Hurst phenomenon, was firstly explored 30 years ago by Klemeš (1974). Klemeš analyzed several variants of the 'changing mean' mechanism which assumes that the mean of the process is not a constant determined by the arithmetic mean of the record, but varies through time. Specifically, he performed numerical experiments with Gaussian random processes that have mean changing according to either a deterministic or a random rule and he found that the composite process behaved Hurst-like. Here, it should be mentioned that Klemeš referred to all his 'changing mean' models as models with nonstationarity in their mean, even though this is strictly true only for the models with deterministic change of the mean. Even though he kept the term 'nonstationary' for all changes, he noted that some of his models were in fact stationary. However, this note has sometimes been missed and led to a misconception about his work by some authors (regrettably, including this one).

More recently, in Koutsoyiannis (2002) this thinking was extended on the setting of multiple timescales of fluctuation of mean, also giving emphasis on the stationarity of the composite process. Specifically, it was demonstrated that a Markovian process, the simplest process that implies short-range dependence and obeys the classical statistical law (3), becomes virtually indistinguishable from an SSS process if there occur fluctuations of its mean at least at two time scales (e.g. one of the order of 10 years and one of the order of 100 years). The fluctuations were assumed random and stationary, again with Markovian structure, which implies that the final model, i.e. the superimposition of the three individual models, is stationary.

3.4 The uncertainty under the scaling hypothesis

Contrary to the nonstationarity approach, which as demonstrated in subsection 2.6 implies that a behaviour characterized by trends decreases uncertainty, the scaling approach shows that the uncertainty increases dramatically due to this behaviour. This is known to engineering hydrologists since the discovery of the Hurst phenomenon, as the discovery was related to the design of reservoirs and the Hurst behaviour implies larger storages of reservoirs (see e.g. Klemeš et al., 1981). Perhaps it is not equally known to scientific hydrologists and mainly to climatologists who insist using classical statistics more than half a century after the discovery of the scaling behaviour, even though the most impressive effects of this behaviour emerge at long time scales that are used in climatic studies.

In brief, there are two factors that increase the uncertainty under the scaling behaviour. The first applies to all time scales and is related to the bias of the classical statistical estimator of standard deviation in the case of an SSS process. The second, which is more important, applies to time scales higher than 1 (the basic time scale) and is related directly to the modified law (4). In our Boeoticos Kephisos runoff example, the classical estimate of standard deviation of runoff at the annual scale, estimated for the entire 96-year series, is 69.4 mm. A modified SSS estimator (Koutsoyiannis, 2003) results in a value 74.5 mm, i.e. in an increase of variance by 15%, which may be not very important. The difference becomes more important if we go to larger scales. Let us consider for instance the scale of 30 years, which is

typically regarded in climatology as a time scale sufficient to smooth out transient characteristics of a time series and yield a value representative of the climate. Also, let us assume for simplicity that the values 69.4 mm and 74.5 mm, estimated by the two approaches, are the true standard deviations for each of the two approaches (i.e., there is no error in estimation, which, however, is not correct, thus resulting in underestimation of uncertainty). The classical statistical law (3) implies a standard deviation of the climatic variable $\bar{X}_{30}$, $\text{StD}[\bar{X}_{30}] = 69.4/30^{0.5} = 12.7$ mm. The corresponding value of the SSS law (4) is $\text{StD}[\bar{X}_{30}] = 74.5/30^{1-0.79} = 36.5$ mm.

Now if we had adopted the nonstationarity approach, according to which the annual standard deviation was reduced to 56.3 mm (see section 2.6), then $\text{StD}[\bar{X}_{30}]$ would be $56.3/30^{0.5} = 10.3$ mm. This shows that the climatic uncertainties for the time scale of 30 years, as they result from the nonstationarity and the scaling modelling approaches differ by a factor of about 3.5! The differences would be even higher if the uncertainty due to parameter estimation were also considered. Furthermore, in the scaling approach the variability of the climatic variable $\bar{X}_{30}$ is no less than half the annual variability. This is consistent with the perception of a changing climate and simultaneously it does not require any additional mechanism, either deterministic or random, to explain the change.

As already discussed, the quantification of uncertainty is not just a theoretical issue of general scientific interest, but it affects seriously the design and management of hydrosystems. The well known impact to the design of large reservoirs was already mentioned in the beginning of this section. To emphasize the impact to water resource management we will refer to a recent experience from the Athens water supply system. This is a huge system extending on an area of about 4000 km^2 and comprising four reservoirs and several aquifers. One of the reservoirs, the natural Lake Hylike, is fed by the Boeoticos Kephisos River. Recent attempts to rationalize the management of the system, initially were based on the classical statistical approach, according to which the decreases in rainfall and runoff (Figure 2(b) and (c)) were regarded "deterministic trends" (e.g. Nalbantis et al., 1993). However, a persistent (7 years) drought during 1987-93 for the entire system (seen in Figure 2(b) for Boeoticos Kephisos), which shook Athens, could be hardly described (reproduced) by

the classical approach, whose uncertainty estimates were too narrow as discussed above. This triggered the shift to the scaling modelling approach, which can easily reproduce persistent droughts. Soon it was discovered that the scaling approach could also reproduce changing means, so there was no need to consider any "external" deterministic trend. Based on this approach, a decision support system was constructed (Koutsoyiannis et al., 2003), which is in use since 2000 and operates in satisfactory manner (as verified for instance by the good performance in achieving water supply adequacy for the Olympic year 2004; Koutsoyiannis, 2004).

The above presentation of uncertainty estimation according to the scaling approach was deliberately made simple, in accordance to the scope of this paper. In an operational application, such as in the management of the Athens water resource system, the uncertainty estimation and its implementation to management decisions is more complicated. In this case, the scaling approach, far from ignoring the series of past observations (which for instance may indicate a local trend), does full exploit them to estimate uncertainty conditional on the observed past so as to support decisions for future. The method of choice to accomplish this is Monte Carlo simulation. A conditional simulation methodology (appropriate for a process with scaling behaviour), according to which all known history of the process is incorporated in the generation of future trajectories, has been described in detail elsewhere (Koutsoyiannis, 2000).

4. Conclusion

The perception of a changing climate, which impacts also hydrological processes, is now generally admitted. However, the way of handling the changing nature of climate in hydrologic practice and especially in hydrological statistics has not become clear so far. The most common modelling approach is to assume that long-term trends, which have been found to be omnipresent in long hydrological time series, are "deterministic" components of the time series and the processes represented by the time series are nonstationary. In this paper, we attempted to show that this approach is contradictory in its rationale and even in the terminology it uses. As a result, it may imply misleading perception of phenomena and

estimates of uncertainty. Besides, it was maintained that a stochastic modelling approach hypothesizing stationarity and simultaneously admitting a scaling behaviour reproduces climatic trends (considering them as large-scale fluctuations) in a manner that is logically consistent, easy to apply and free of paradoxical results about uncertainty. In fact the scaling approach implies that the uncertainty is greatly increased due to large-scale fluctuations.

Obviously, the scaling approach is a modelling approach, and no model is perfect. Its stationary character is sufficient to describe large-scale natural climatic fluctuations but it may not be appropriate for cases where there is evidence of causative effects of a certain event on some phenomenon. For example, a stationary process is not appropriate to model runoff in a basin with known large scale changes of the natural hydrosystem (e.g. construction of dams) or the land use in the basin (e.g. deforestation, urbanization, elimination of floodplain by dikes). In such cases, a scaling stochastic model should be coupled with a deterministic hydrological model that will be able to capture the deterministic mechanisms of changes (not based on statistical analyses). Another case where the stationary approach may be insufficient is related to anthropogenic climate change scenarios, as these may imply a change that is not already reflected in past data. Theoretically, it is possible to adapt the scaling approach adding a nonstationary component, whose mathematical form should be provided by deterministic climatic models. But this will have meaning only when climatic models will be able to establish reliable relationships between hydro-climatic processes and factors affecting them and explain the already verified scaling behaviour of the past. Until then, a good step to a more faithful description of uncertainty is to adapt the classical statistical thinking and implementation towards exploiting the scaling hypothesis, which is more consistent to what is observed in nature.

References

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

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

Brock, W. A., and Potter, S. M. (1992). Diagnostic testing for nonlinearity, chaos, and general dependence in time series data. In *Nonlinear Modeling and Forecasting*, edited by Casdagli M. and Eubank S., pp. 137-161. Addison-Wesley, Redwood City.

Çengel, Y. A., and Boles, M. A. (2002). *Thermodynamics, An Engineering Approach*, 4th edition, McGraw-Hill, Boston

Clarke, R. T. (1998). *Stochastic Processes for Water Scientists*, Wiley, Chichester.

Chow, V. T. (1964). Statistical and probability analysis of hydrology data. Part I. Frequency Analysis. In *Handbook of Applied Hydrology*, edited by Chow V. T., Section 8-I, pp. 8.1-8.42, McGraw-Hill, New York.

Dawdy, D. R., and Matalas, N. C. (1964): Analysis of variance, covariance and time series. In *Handbook of applied Hydrology*, edited by Chow V. T., Section 8-IIII, pp. 8.68-8.90, McGraw-Hill, New York.

Eagleson, P. S. (1970). *Dynamic Hydrology*, McGaw-Hill, NY 1970.

Flato, G. M. and Boer, G. J. (2001). Warming asymmetry in climate change simulations, *Geophys. Res. Lett.*, 28, 195-198

Freund, J. E., Williams, F. J., Perles, B. M., (1988). *Elementary business statistics*, 5th edition, Prentice Hall, Englewood Cliffs, New Jersey.

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

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

Karlin, S., and Taylor, H. M., (1975). A First Course in Stochastic Processes, 2nd edition, Academic Press, San Diego.

Klemeš, V. (1974). The Hurst phenomenon - a puzzle?", *Water Resour. Res.*, 10(4), 675-688.

Klemeš, V. (2002). Geophysical time series and climatic change − A sceptic's view. In *Hydrological Models for Environmental Management*, edited by Bolgov M.V. et al., Canadian Water Resources Association.

Klemeš, V., Sricanthan, R. and McMahon T. A. (1981). Long-memory flow models in reservoir analysis: What is their practical value? *Water Resour. Res.*, l7(3), 737-75l.

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

Kotz S. and Johnson N. L. (editors) (1988). *Encyclopedia of Statistical Sciences*. Entry: Time series, nonstationary, Vol. 9, p. 257, Wiley, New York.

Koutsoyiannis, D. (2000), A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research*, 36(6), 1519-1533.

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

Koutsoyiannis, D. (2003). Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, 48(1), 3-24.

Koutsoyiannis, D. (2004), The water resource management of Athens in the perspective of the Olympic Games, *The Olympic Games Athens 2004 and the National Technical University of Athens*, edited by K. Moutzouris, pp. 17-27, National Technical University of Athens, Athens (in Greek).

Koutsoyiannis, D. (2005a). Uncertainty, entropy, scaling and hydrological stochastics, 2, Time dependence of hydrological processes and time scaling, *Hydrological Sciences Journal*, 50(3), 405-426.

Koutsoyiannis, D. (2005b). A toy model of climatic variability with scaling behaviour, *J. Hydrol.* (in press).

Koutsoyiannis, D., and Efstratiadis, A., (2004). Climate change certainty versus climate uncertainty and inferences in hydrological studies and water resources management, *1st*

General Assembly of the European Geosciences Union, Geophysical Research Abstracts, Vol. 6, Nice (http://www.itia.ntua.gr/g/docinfo/606/).

Koutsoyiannis, D., Karavokiros, G., Efstratiadis, A., Mamassis, N., Koukouvinos, A., and Christofides, A. (2003), A decision support system for the management of the water resource system of Athens, *Physics and Chemistry of the Earth*, 28(14-15), 599-609.

Lasota, A., and Mackey, M. C. (1994). *Chaos, Fractals, and Noise*, Springer-Verlag, New York.

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

Xanthopoulos, T., Prédictions hydrologiques et incertitudes banales, Annales de l'Académie des Sciences, Paris, 1974.

Papoulis, A. (1991). *Probability, Random Variables, and Stochastic Processes*, 3rd ed., McGraw-Hill, New York.

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

Rozos, E., Efstratiadis, A., Nalbantis, I., and Koutsoyiannis, D. (2004). Calibration of a semi-distributed model for conjunctive simulation of surface and groundwater flows, *Hydrological Sciences Journal* , 49(5), 819-842.

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

Salas, J. D., Delleur, J. W., Yevjevich, V., and Lane, W. L. (1980). *Applied Modeling of Hydrologic Time Series*, Water Resources Publications, Littleton, Colo.

Taylor, H. M., and Karlin, S., (1984). *An Indroduction to Stochastic Modeling*, Academic Press, Boston.

Theiler, J., Linsay, P. S., and Rubin, D. M. (1994). Detecting nonlinearity in data with long coherence times. In *Time Series Prediction*, edited by Weigend, A. S., and Gershenfeld, N. A., pp. 429-456, Addison-Wesley, Reading, Massachusetts.

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

von Storch, H. (1995) Misuses of statistical analysis in climate research. In: *Analysis of Climate Variability: Applications of Statistical Techniques*, edited by von Storch H. and Navarra, A., pp. 11–26. Springer-Verlag, Berlin.

Weisstein, E. W. (1999-2005a). Nonstationary time series. From MathWorld-A Wolfram Web Resource. http://mathworld.wolfram.com/NonstationaryTimeSeries.html.

Weisstein, E. W. (1999-2005b). Stochastic Process. From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/StochasticProcess.html

Yevjevich, V. (1972). *Stochastic Processes in Hydrology*, Water Resources Publications, Fort Collins, Colorado.