Are hydrologic processes chaotic?

Demetris Koutsoyiannis
Department of Water Resources, Faculty of Civil Engineering,
National Technical University, Athens
Heroon Polytechneiou 5, GR-157 80 Zographou, Greece
(dk@hydro.ntua.gr; http://www.hydro.civil.ntua.gr/faculty/dk/)

Paper submitted to Water Resources Research

August 2001

Abstract.

Several recent studies have claimed for discovering low-dimensional determinism in hydrologic processes, such as rainfall and runoff, using methods of chaotic analysis. Such results, however, are questionable. It is shown that in some cases merely the careful application of concepts of dynamical systems, without doing any calculation, provides strong indications that hydrologic processes cannot be (low-dimensional) deterministic chaotic. Furthermore, it is shown that specific peculiarities of hydrologic processes on fine scales, such as asymmetric, J-shaped distribution functions, intermittency, and high autocorrelations, are synergistic factors that can lead to misleading conclusions regarding presence of (low-dimensional) deterministic chaos. In addition it is shown that to accurately estimate chaotic descriptors of hydrologic processes huge data sets are demanded; the required size is quantified by statistical reasoning and is not met in hydrologic records. All these arguments are demonstrated using appropriately synthesized theoretical examples. Finally, in light of the theoretical analyses and arguments, typical real-world hydrometeorological time series, such as relative humidity, rainfall, and runoff, are explored and none of them is found to indicate the presence of chaos but, rather, all must be regarded as the outcomes of stochastic systems.

GAP index terms: 1869 Stochastic processes, 3240 Chaos, 3220 Nonlinear dynamics

1. Introduction

The impressive results of chaos analysis of simple physical and mathematical systems in the last two decades offered an alternative way to view natural systems. Specifically, it became clear that a simple deterministic system, even with one degree of freedom, can have a complex, random-appearing evolution. Obviously, however, the inverse is not true: Complex or erratic-appearing phenomena can be chaotic and thus deterministic, but they can also well be random.

Classifying a physical system either as a random one or as a chaotic deterministic one is not an easy task. For example, the throw of dice, still (many years after Albert Einstein's famous apothegm) is regarded as the most typical example of a random system, even though its outcome depends on a few collisions of a cube onto a plane, whose deterministic dynamics can be understood rather easily. On the other hand, more recently, much more complex natural processes, like meteorological and hydrological, have been regarded by many researchers as chaotic systems.

Loosely speaking, the complexity of a deterministic system depends on the number of degrees of freedom, or dimension of the system attractor, and on how many of them are associated with sensitive dependence on initial conditions. The latter are quantified by the positive Lyapunov exponents associated to the system dynamics. Chaotic systems are in fact the simplest possible deterministic systems with sensitivity to initial conditions: those that they have one positive Lyapunov exponent [*Kantz and Schreiber*, 1997, pp. 183, 241], and typically have attractor dimension less than two [*Kantz and Schreiber*, 1997, p. 183]. For systems with more than one positive Lyapunof exponent the term *hyperchaos* has been coined [*Rössler*, 1979; *Kantz and Schreiber*, 1997, pp. 183, 241]. Alternatively, chaotic and hyperchaotic systems are also known as low-dimensional chaotic and high-dimensional chaotic, respectively. Following *Kantz and Schreiber*[1997, pp. 183, 241], in this paper the term low-dimensional chaos is used as synonymous to chaos.

While numerous chaotic systems have been studied thoroughly, only few experimental observations of hyperchaos have been recorded. To explain this lack of higher dimensional experimental attractors, *Kantz and Schreiber* [1997, p. 241] offer two possible explanations: maybe typical systems in nature possess either exactly one or very many positive Lyapunof exponents, or the reason is that systems with a higher than three dimensional attractor are very difficult to analyze. The systems with very many positive Lyapunof exponents are better modeled based on stochastic models. Theoretically, stochastic models imply infinitely many degrees of freedom (infinite dimensional systems).

Traditionally, stochastic models have been the preferred mathematical tools in hydrology and water resources modeling. Hydrologic processes have been most frequently modeled as stochastic processes, which can easily incorporate any existing deterministic component of the natural processes (e.g., periodicity) in addition to random components. However, in the last decade, the charming possibility that a complex hydrologic system with irregular time evolution may au fond be a simple chaotic system has motivated several researchers to analyze hydrologic processes using mathematical tools of the chaos literature. Their intention and hope was to discover simplicity and universal determinism in place of what was earlier considered as weak deterministic components superimposed on random components. Thus, an increasing number of studies have tried to show that hydrologic processes are chaotic. *Sivakumar* [2000] reviews most of the studies related to chaotic analysis of hydrologic processes. Table 1 summarizes a number of such studies with their most important characteristics and findings. The processes considered are mainly rainfall and runoff, and in one case (#9) lake storage. The time scales used vary from 15 seconds to one month. In two cases (#3 and #5) the data series used are not regular time series but inverse time series (raingauge tip time corresponding to increase of rainfall depth by 0.01 mm). The data sizes vary from 1572 to 70 000. Only in three cases (#1, #4, and #8) the analyses indicate absence of determinism, whereas in 12 cases the authors claimed that they discovered deterministic attractors with dimensions varying from 0.45 to 9.

The attempts to discover chaos in natural phenomena are not exhausted to hydrology. As pointed out by *Provenzale et al* [1992],

> "… the desire for finding a chaotic attractor has led to a naïve application of the analysis methods; as a result, the number of claims on the presence of strange attractors in vastly different physical, chemical, biological and astronomical systems has grown (exponentially?)".

Here they quote a statement by *Grassbrerger et al.* [1991]:

> "… most (if not all) of these claims have to be taken with much caution".

They also note that convincing evidence for chaos most commonly arises when spatial complexity of the system is limited, a condition that could be true for experimental systems, but is by far untrue for hydrologic and other geophysical processes.

These quotes have not been the only ones expressing skepticism about the discovery of chaos in natural phenomena. *Ghilardi and Rosso* [1990] in their discussion of the work by *Rodriguez-Iturbe et al.* [1989] (item #2 in Table 1) provide several arguments to show that the latter work gives insufficient support to the presence of chaotic dynamics with a strange attractor. *Wilcox et al.* [1991] (item #4 in Table 1) point out that runoff is complex not because of underlying low-dimensional chaotic dynamics, but rather, because of complicated interactions of the many factors that affect runoff. *Koutsoyiannis and Pachakis* [1996] (item #8 in Table 1) show that an observed rainfall time series and a synthetic one generated by a well structured stochastic model are indistinguishable from each other even if tools of chaotic dynamics are used to characterize and compare the series; thus, given that the stochastic series is infinite dimensional, the observed series must be infinite dimensional, too.

The present paper attempts to proceed a step further than simply express skepticism about the discovery of chaos in hydrologic processes. Specifically, it shows that the hypothesis that hydrologic time series manifest stochastic, rather than chaotic, systems cannot be rejected using the standard procedures of chaotic analysis. In addition, it locates critical points that may lead to an erroneous conclusion that a stochastic hydrologic system is chaotic; such issues may have influenced earlier studies that identified chaos in hydrology. Furthermore, it suggests ways to bypass these critical points and avoid erroneous conclusions.

To this aim, the paper first briefly reviews some basic concepts chaotic behavior, such as concepts of dynamical systems and attractors, delay embedding and reconstruction of dynamics, and the typical procedure for identifying chaos based on the estimation of attractor dimensions (section 2). Subsequently, it shows that in some cases merely the careful application of the concepts of dynamical systems, without doing any calculation, provides strong indications that hydrologic processes cannot be chaotic. Furthermore, it shows that specific peculiarities of hydrologic processes on fine scales, such as asymmetric, J-shaped distribution functions, intermittency, and high autocorrelations, are synergistic factors that can lead to misleading conclusions regarding presence of chaos, and in addition demand huge data sets, whose size is quantified by statistical reasoning, to accurately estimate chaotic descriptors. All these arguments are demonstrated using appropriately synthesized theoretical examples (section 3). Finally, in light of the theoretical analyses and arguments, typical real-world hydrometeorological time series, such as relative humidity, rainfall, and runoff, are explored and none of them is found to indicate the presence of chaos (section 4). Some mathematical derivations that support the theoretical analysis have been separated from the text and are given in Appendices.

2. Descriptors of chaotic behavior

2.1 Dynamical systems and attractors

The nonlinear time series methods which are applied in hydrology are based on the theory of dynamical systems; these are characterized by (a) a phase or state space on which the motion of the system takes place, (b) a rule stating where to go next from the system current position (also known as system dynamics), and (c) a time set that describes the moments at which movements from one position to another take place.

Typically, the phase space M is a finite-dimensional vector space $\mathbb{R}^m$ and the state of the system is specified by a vector $\mathbf{x}$ with size m. The time set is typically either the set of integers $\mathbb{Z}$ (discrete time) or the set of real numbers $\mathbb{R}$ (continuous time). The system dynamics is a family of transformations $\mathbf{S}_t$: $M \to M$ (where t denotes time) satisfying [*Lasota and Mackey*, 1994, p. 191]

$$\mathbf{S}_0(\mathbf{x}) = \mathbf{x}, \quad \mathbf{S}_t(\mathbf{S}_{t'}(\mathbf{x})) = \mathbf{S}_{t+t'}(\mathbf{x}), \quad \mathbf{x} \in M \tag{1}$$

In discrete time, the system dynamics is completely determined by the m-dimensional map $\mathbf{S}_1$, i.e.,

$$\mathbf{x}_{n+1} = \mathbf{S}_1(\mathbf{x}_n), \qquad n \in \mathbb{Z} \tag{2}$$

In continuous time the dynamics is described as a system of m ordinary differential equations

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{s}(\mathbf{x}(t)), \qquad t \in \mathbb{R} \tag{3}$$

whose solution defines the family of transformations $\mathbf{S}_t$.

For a given initial point $\mathbf{x}_0$ or $\mathbf{x}(0)$ the sequence of points $\mathbf{x}_n = \mathbf{S}_n(\mathbf{x}_0)$ or the function $\mathbf{x}(t) = \mathbf{S}_t(\mathbf{x}(0))$ considered as a function of n or t is called a trajectory of the dynamical system. In the so called dissipative dynamical systems, characterized by $|\det \mathbf{J}| < 1$ (where $\mathbf{J} := d\mathbf{S}_1 / d\mathbf{x}$ is the Jacobian matrix of $\mathbf{S}_1$) or div $\mathbf{s} < 0$, the trajectory of the system, after some transient time, is attracted to some subset A of the phase space. This set itself is invariant under the dynamical evolution ($\mathbf{S}_t(A) = A$) and is called the attractor of the system [*Kantz and Schreiber*, 1997, p. 32]. Only three types of attractors can occur [e.g., *Lasota and Mackey*, 1994, p. 192; *Kantz and Schreiber*, 1997, p. 32]: (a) fixed points indicating that the system settles to a stagnant state, i.e.,

$$\mathbf{x}_n = \mathbf{x}_0 \quad \text{or} \quad \mathbf{S}_t(\mathbf{x}(0)) = \mathbf{x}(0), \qquad \text{for all } n \text{ or } t \tag{4}$$

(b) limit cycles, indicating periodic motion with period ω, i.e.,

$$\mathbf{x}_{n+\omega} = \mathbf{x}_0 \quad \text{or} \quad \mathbf{S}_{t+\omega}(\mathbf{x}(0)) = \mathbf{x}(0), \qquad \text{for all } n \text{ or } t \tag{5}$$

and (c) nonintersecting trajectories, in which case

$$\mathbf{x}_{n_1} \neq \mathbf{x}_{n_2} \quad \text{or} \quad \mathbf{S}_{t_1}(\mathbf{x}(0)) \neq \mathbf{S}_{t_2}(\mathbf{x}(0)), \qquad \text{for all } n_1 \neq n_2 \text{ or } t_1 \neq t_2 \tag{6}$$

For a system in continuous time with a 2-dimensional state space the fixed point and cycle are the only possibilities, whereas for 3 dimensions and beyond the more interesting nonintersecting attractors can occur, which typically exhibit fractal structure and are called strange attractors. For systems in discrete time the nonintersecting attractors can occur even in a 2-dimensional state space.

2.2 Delay embedding and reconstruction of dynamics

In this paper, as in other similar hydrologic applications of chaotic dynamics, we consider only systems expressed in terms of a single scalar real quantity y (e.g., rainfall, runoff, etc.). Such a system evolves in continuous time, and its m-dimensional state $\mathbf{x}$ is theoretically expressed in terms of the quantity y and a number $m - 1$ of its derivatives with respect to time, i.e., $\mathbf{x}(t) = [y(t), y'(t), \ldots\, y^{(m-1)}(t)]^T$ (where ($y^{(k)} = d^k y \,/\, dy^k$ and the superscript T denotes the transpose of a vector or matrix).

However, in practice we can only have observations of the quantity y on discrete time intervals Δt and no observations of its derivatives at all. Therefore, we study the system as if it were a discrete time system using the so-called delay vectors

$$\mathbf{x}_n := [y_n, y_{n-\tau}, \ldots, y_{n-(m-1)\tau}]^T \qquad (7)$$

where we denote $y_n = y(n\,\Delta t)$ and τ is a positive integer. By studying the simplified discrete time system we can infer the properties of the original system since, according to Takens' embedding theorem [*Takens*, 1981], for properly chosen embedding dimension m and time delay τ, the discrete time system will trace out a trajectory that represents a smooth coordinate transformation of the original trajectory of the system.

Thus, the Takens theorem allows for the reconstruction of the dynamics of the system using a time series of a single scalar observable. If the only given information is the time series, in the beginning we do not know what is the proper embedding dimension m. This dimension depends on the dimension D of the attractor. The latter dimension has important content as it represents the number of local directions available to the system and so it provides an estimate of the number of degrees of freedom needed to describe the state of the system [*Gershenfeld and Weigend*, 1993, p. 48].

According to *Whitney's* [1936] embedding theorem, which was generalized for fractal objects by *Sauer et al.* [1991], any D-dimensional object (precisely, any D-dimensional smooth manifold) can be embedded in an m-dimensional Euclidean space if $m > 2D$. For example, a one-dimensional curve of any shape can always be embedded in a 3-dimensional Euclidean space (and all higher-dimensional spaces), but it cannot be embedded in a 2-dimesional space because, except for special cases, it will overlap itself (this will be further clarified in section 3.1; see also *Kantz and Schreiber* [1997, p. 128]). Thus, an attractor of the nonintersecting type with dimension 1 will intersect itself in a 2-dimesional space but not in a 3-dimesional space.

Therefore, if we knew the dimension D of the attractor, we would choose the state vector size m as the smallest integer that is greater than $2D$. But since we do not know D when we work with merely the time series, we follow an iterative procedure. For trial $m = 1, 2, \ldots$, we estimate the dimension $D(m)$ of the trajectory of the system at the m-dimensional space, until $D(m)$ becomes constant with the further increase of m. This constant value is the dimension of the attractor.

2.3 Estimation of dimensions

The problem arises then how we can estimate the dimension D of a trajectory or attractor A in an m-dimensional vector space. The estimate of a dimension is typically done in terms of either entropies or correlation sums. Here the description of entropies and correlation sums is generalized for any set A that is a subset of an m-dimensional metric space with a normalized measure $\mu(\,)$ defined on its Borel field. In our case, for $m = 1$, the set A may represent all possible values of a hydrological variable such as rainfall or runoff at a specified time scale, which is the set of positive real numbers $\mathcal{R}^{+}$. It may also represent all values in a certain observed time series of the same variable, in which case A is a finite subset of R^{+}. Accordingly, for $m > 1$, the set may represent the delay vectors.

Let us consider a partition of the set A into $v(\varepsilon)$ boxes (hypercubes) $A_1, A_2, \ldots, A_{v(\varepsilon)}$ with length scale (i.e., edge length of each hypercube) ε. The generalized entropy of order q of A, denoted as $I_q(\varepsilon)$ is [*Rényi*, 1970]

$$I_q(\varepsilon) := \frac{1}{1-q} \ln \sum_{i=1}^{v(\varepsilon)} p_i^q \tag{8}$$

where p_i is the measure of the part of the set A contained in the ith hypercube, that is, $p_i = \mu(A_i)$, such that

$$\sum_{i=1}^{v(\varepsilon)} p_i = 1 \tag{9}$$

If the set A consists of N observed values (points in the m-dimensional space) and N_i of them are contained in the ith hypercube A_i, then $p_i = N_i / N$. Accordingly, if this set is the sample space of a vector of random variables $\mathbf{X}$ then each hypercube A_i represents an event and $p_i = \Pr(\mathbf{X} \in A_i)$ where Pr( ) denotes probability.

Definition (8) applies for $q \neq 1$. Taking the limit for $q \to 1$ and using de l'Hospital's rule we get

$$I_1(\varepsilon) := -\sum_{i=1}^{v(\varepsilon)} p_i \ln p_i \tag{10}$$

which is the typical definition of entropy in probability theory.

The generalized dimension of order q of the set under examination is defined by the following equation [*Grassberger,* 1983]:

$$D_q = \lim_{\varepsilon \to 0} \frac{-I_q(\varepsilon)}{\ln \varepsilon} \tag{11}$$

Applying de l'Hospital's rule in (11) we get

$$D_q = \lim_{\varepsilon \to 0} \frac{d(-I_q(\varepsilon))}{d(\ln \varepsilon)} \tag{12}$$

The latter expression is more advantageous than (11) for numerical applications since the convergence of the derivative is faster.

For low values of q we have the most frequently used dimensions. Thus, for $q = 0$ we have the so-called box counting or capacity dimension D_0, for $q = 1$ the information dimension D_1, and for $q = 2$ we have the correlation dimension D_2. For simple geometrical objects such as lines or surfaces, all D_q are equal to the integer topological dimension (1 for a line, 2 for a surface, etc.). For more complex mathematical objects including (but not exclusively) fractal objects, they are not necessarily integers, nor all D_q are necessarily equal to each other. The most important among generalized dimensions is the capacity dimension D_0, because this is in fact the one used in the extension by *Sauer et al.* [1991] of the *Whitney's* [1936] embedding theorem mentioned above. However the most frequently used is the correlation dimension D_2, because it is more accurately estimated via the so-called correlation sums or integrals.

For a finite sample A of observations **x**, the generalized correlation sum of order q, for integer $q \geq 2$, is defined by

$$C_q(\varepsilon) = N^{-q}\ \{\text{number of } q\text{-tuples } (\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_q}) \text{ with all } \|\mathbf{x}_{j_s} - \mathbf{x}_{j_r}\| < \varepsilon\} \tag{13}$$

This was introduced by *Grassberger* [1983] and has the important property

$$C_q(\varepsilon) \approx \exp[(1-q)\, I_q(\varepsilon)] \tag{14}$$

Thus, for integer $q \geq 2$, we can replace $-I_q(\varepsilon)$ with $\ln C_q(\varepsilon) / (q-1)$ in the calculation of dimensions using the above equations, since the estimation of $C_q(\varepsilon)$ is more accurate than that of $I_q(\varepsilon)$ [*Grassberger and Procaccia*, 1983; *Grassberger*, 1983]. In practice however, only the correlation sum $C_2(\varepsilon)$ for $q = 2$ is used, because the calculation of higher-order sums is extremely time consuming. (In fact, even for $q = 2$ the calculation is time consuming. The correlation sum of order 2, or simply correlation sum, is given by the following equation that is a consequence of (13):

$$C_2(\varepsilon) = \frac{2}{(N-k)\,(N-k-1)} \sum_{i=1}^{N} \sum_{j=i+w}^{N} H(\varepsilon - \|\mathbf{X}_i - \mathbf{X}_j\|) \tag{15}$$

where H is the Heaviside's step function, with $H(u) = 1$ for $u > 0$ and $H(u) = 0$ for $u \leq 0$ and w an integer constant, which for uncorrelated time series is assumed zero but for correlated ones takes a nonzero value to exclude from the estimation those pairs of points that are close in time [*Kantz and Schreiber*, 1997, p. 74]. For the calculation of the distance $\|\mathbf{X}_i - \mathbf{X}_j\|$, the maximum norm is usually used as it reduces the computational time [*Hübner et al.,* 1993].

The correlation sum $C_2(\varepsilon)$ expresses the average proportion of pairs of points having distance smaller than ε between them.

2.4 Typical procedure for identifying chaos

Typically, the estimation of the correlation dimension D_2 of a set A of delay vectors $\mathbf{x}$ in an m-dimensional space, constructed from a series of N observations, is based on the correlation sum. The estimation procedure, known as Grassberger-Procaccia algorithm (after *Grassberger and Procacia* [1983]) consists of the following typical steps:

1. Calculate the correlation sum $C_2(\varepsilon, m)$ for several values of the length scale ε.

2. Make a log-log plot of $C_2(\varepsilon, m)$ versus ε and a plot of the local slope $d_2(\varepsilon, m)$versus log ε, where

$$d_2(\varepsilon, m) := \frac{\Delta[\ln C_2(\varepsilon, m)]}{\Delta[\ln \varepsilon]} \tag{16}$$

and locate a region with constant slope, known as a scaling region [e.g., *Hübner et al.*, 1993].

3. Calculate the slope of the scaling region, which is the estimate of the correlation dimension $D_2(m)$ of the set for the embedding dimension m.

As explained above this is done iteratively for $m = 1, 2, \ldots$ and iterations stop when $D_2(m)$ saturates to a constant value D_2, independent of m. The convergence of $D_2(m)$ to the value D_2 verifies that a D_2-dimentional attractor (a) exists, which means that the system under study is deterministic; (b) has been identified; and (c) can been embedded in an m-dimensional space where m is the minimum integer for which $D_2(m) = D_2$. Conversely, if $D_2(m)$ does not become constant for increasing m the system is characterized as stochastic, rather than deterministic. Thus, the above procedure has been used to identify whether a system is chaotic deterministic or stochastic, using only a time series of an observable of the system. This procedure has been applied in most of the hydrologic applications mentioned in the introduction.

Several authors have warned that the procedure has several critical points that must receive a lot of attention [see discussions in *Tsonis*, 1992; *Tsonis et al.*, 1993; *Kantz and Schreiber*, 1997; *Graf von Hardenberg*, 1997a; *Sivakumar*, 2000; among others], otherwise the results may be flawed. In the next section 3 we will show that such critical points may have not given the required attention in some hydrologic applications that claim for chaos in hydrologic processes. We also introduce some additional critical points whose ignorance could result in erroneous interpretations.

3. Important issues in identifying chaos in hydrological processes

3.1 A conceptual approach to the dimensionality of a hydrologic attractor

Before applying any algorithm to quantify the dimensionality of an attractor in a hydrological process, it would be a good idea to try a more conceptual approach and to determine, if possible, what would be a reasonable expectation of this dimensionality.

Let us start with the rainfall process in discrete time on daily scale (the same reasoning applies in finer timescales as well). For this process and scale we observe in Table 1 that two studies (items #6 and #12) have claimed for chaos with dimensionality D_2 as low as 1 to 2.5, whereas other two studies have argued that no saturation appears (items #4 and #8), which may mean that the rainfall process is better modeled as a stochastic process (with infinite dimensionality).

In a daily rainfall time series there exist periods with zero rainfall. Let k be the maximum observed dry period in days. For example, in Athens, Greece, from a 132-year record of rainfall record we know that k = 120 days (4 months). We set n = 1 the day when this dry period starts, so that the rainfall depths y_n for n = 1 to k are all zero. Let us assume that the rainfall at the examined location is the outcome of a deterministic system whose attractor can be embedded in $\mathcal{R}^m$ for some integer m. This attractor is reconstructing using delay embedding with delay τ. Furthermore, let us assume that $m < (k - 1) / \tau + 1$. Then, there exist at least two delay vectors with all their components equal to zero. Namely, $\mathbf{x}_k = [y_k, y_{k-\tau}, y_{k-2\tau}, \ldots, y_{k-(m-1)\tau}]^T = \mathbf{0}$ and $\mathbf{x}_{k-1} = \mathbf{x}_k = [y_{k-1}, y_{k-1-\tau}, y_{k-1-2\tau}, \ldots, y_{k-1-(m-1)\tau}]^T = \mathbf{0}$ where $\mathbf{0}$ is the zero vector. Therefore $\mathbf{x}_k = \mathbf{S}_1(\mathbf{x}_{k-1}) = \mathbf{S}_1(\mathbf{0}) = \mathbf{0}$, and since the system is deterministic, it will result in $\mathbf{x}_n = \mathbf{0}$ for any $n > 0$ (since $\mathbf{x}_{k+1} = \mathbf{S}_1(\mathbf{x}_k) = \mathbf{S}_1(\mathbf{0}) = \mathbf{0}$, etc.). That is, given that rainfall is zero for a period k, it will be zero forever, which means that the attractor is a single point. This of course is absurd and thus the embedding dimension should be $m \geq (k-1)/\tau + 1$. Now, *Whitney's* [1936] embedding theorem [*Kantz and Schreiber*, 1997, p. 126] tells us that the attractor should have dimension $D \geq (m-1)/2$ and, hence, $D \geq (k-1)/2\tau$. For example, if the maximum dry period is four months as in the Athens example ($k = 120$) and we assume a 'safe' delay $\tau = 10$ (we will discuss it further in section 3.5), the above analysis results in an embedding dimension at least 12 and an attractor dimension at least 6.

This value of the attractor dimension is much higher that the ones shown in Table 1 for the daily rainfall applications. In fact, however, it is still too low. In this reasoning, we have considered rainfall as a discrete time process. If we consider it as a continuous time process, as in fact is, then instead of assuming $\mathbf{x}$ as a vector of delay coordinates, we must regard it as $\mathbf{x}(t) = [y(t), y'(t), \ldots\, y^{(m-1)}(t)]^T$, as explained in section 2.2. Now, at any time within a dry period we have $\mathbf{x}(t) = \mathbf{0}$ regardless of the dimension m used (the rainfall depth and all its derivatives of any order are zero). Clearly, then the attractor cannot be of the nonintersecting type (since $\mathbf{x}(t) = \mathbf{0}$ for several, in fact infinite, values of t) but it will be of the fixed-point type, the fixed point being the zero vector. Of course, this is not true, because at some time the system will depart from the 'attracting' zero point. Thus, the system that is described by the rainfall depth is not deterministic but rather stochastic and thus it does not have a finite dimensional attractor.

This reasoning applies also to finer time scales, as well. On coarser time scales, such as monthly, it may be the case (for wet areas) that the zero rainfall values do not occur. However, if the rainfall process is high- or infinite-dimensional on fine timescales, naturally it will be high- or infinite-dimensional on coarser time scales as well. In addition, since rainfall is the input that mobilizes all other hydrologic processes in a catchment, the number of degrees of freedom of any other hydrologic process (e.g. streamflow) will be at least equal to that of rainfall. Moreover, if rainfall is indeed stochastic, stochastic will be all other processes in the catchment.

Until now our conceptual approach does not use any algorithm at all. After the application of the algorithm, it could be a good idea to examine whether its results are conceptually consistent and meaningful. For example, if the attractor dimension was found to be as low as one or even smaller, as indeed happens in some of the applications contained in Table 1 (items #7, #12, and some cases of #6), then it would have a direct geometrical interpretation. To demonstrate what an attractor with dimension 1 looks like we constructed an example from a system with known chaotic dynamics. We started with the well-known logistic equation $z_n = a\, z_{n-1}\,(1 - z_{n-1})$ with $a = 3.97977$, which has one degree of freedom and thus a one-dimensional attractor. Then, to make the attractor more interesting, we routed z_n through a linear filter to obtain the series $y_n := b_0\, z_n + b_1\, z_{n-1} + b_2\, z_{n-2} + b_3\, z_{n-3} + b_4\, z_{n-4}$ with $b_0 = 1$, $b_1 = 2$, $b_2 = 1.5$, $b_3 = 1$, and $b_0 = 0.5$. Here we did not introduce any additional degree of freedom and thus y_n still has a one-dimensional attractor; this was verified using the Grassberger-Procaccia algorithm. The attractor, constructed graphically using 10 000 points, is shown in Figure 1 in a 2-dimensional (upper panel) and a 3-dimensional (lower panel) space. That the dimension of the attractor is one is obvious in both panels, although the 2-dimensional graph is not appropriate to show the nonintersecting type of the attractor (it intersects itself).

Now if we do the same work with a hydrological series, what we obtain is totally different. In Figure 2 we have plotted in a 2-dimensional (upper panel) and a 3-dimensional (lower panel) space an “attractor” using 10 000 points of a daily rainfall series, which will be discussed further in section 4.1. These graphs are typical for any daily rainfall series. One cannot locate any one-dimensional structure in such graphs. On the contrary, the cloud of points fills all space both in two and three dimensions. Therefore its topological dimension, which is expressed by the capacity dimension D_0, equals the embedding dimension, that is, 2 in the upper panel and 3 in the lower panel. As we show in section 3.2, the correlation dimension of this 2- or 3-dimensional space filling cloud could be 1 or even less, but this is totally irrelevant. What matters is the fact that the cloud of points fills up space and, thus, the capacity dimension equals the embedding dimension.

One may argue that the plots of Figure 2 are in two and three dimensions whereas items #7, #12 and #6 of Table 1 show that the embedding dimension should be at least 10 or more, up to 40. Here is another inconsistency of these results. If the attractor dimension was 1 or less, then, according to Whitney's embedding theorem, a three dimensional embedding space would suffice to embed it. The fact that the estimated embedding dimension in these works is 10-40 simply indicates that the results are flawed.

Another type of suspect results, which we meet in couples of items #6 and #7, and #13 and #14 of Table 1, is the fact that runoff appears to have an attractor with dimension lower than that of rainfall at the same area and timescale. As explained above, it is difficult to imagine how runoff (hydrologic system output) could have dimension smaller than rainfall (hydrologic system input).

3.2 Capacity versus correlation dimension and the effect of an asymmetric distribution function

Wang and Gan [1998] have pointed out that the underlying distribution function plays a role in the estimation of correlation dimension. This they demonstrated using random data series generated from Gamma and Poisson distributions. They argued that the correlation dimension for these data series is underestimated due to a clustering feature, or an "edging effect". In this section we analyze this issue theoretically and we show that small estimated values of correlation dimension should not necessarily be interpreted as underestimated, as in fact can be correct estimates.

It can be easily shown that in random time series the capacity dimension $D_0(m)$ equals the embedding dimension, m, or, in other words, the time-delayed vectors fill up the embedding space. This has been given a key role in identifying chaos in hydrological processes and particularly in the characterization of a process of chaotic rather than stochastic. However, as discussed in sections 2.2 and 2.3, in identifying chaos the correlation dimension $D_2(m)$ rather than the capacity dimension $D_0(m)$ has been typically used. It is the rule that the correlation dimension of a random series $D_2(m)$ equals $D_0(m)$ and therefore the embedding dimension m. However, we show in Appendix 1 that this rule is valid only for square-integrable probability density functions $f(y)$, i.e., those whose square integral over their domain A is finite, i.e.,

$$\int_A f^2(y)\, dy < \infty \tag{17}$$

In addition, we show that in purely random processes following non-symmetric J-shaped distributions, the rule is not valid and $D_2(m)$ is smaller than m. More specifically, we show that in such processes and for embedding dimension $m = 1$,

$$D_2(1) = 2 + 2 \lim_{\varepsilon \to 0} \frac{\varepsilon f'(\varepsilon)}{f(\varepsilon)} < 1 = D_0(1) \tag{18}$$

where $f'(\,)$ is the derivative of $f(\,)$. By analogy, $D_2(m) = m\, D_2(1) < m$.

In addition, in Appendix 1 we show that in distribution functions typically used in hydrology such as Pareto, Gamma and Weibull, with shape parameter κ smaller than 1/2 or, equivalently, coefficients of skewness greater than 0.639, 2.83 and 6.62, respectively, the correlation dimension for embedding dimension 1 is $D_2(1) = 2\,\kappa < 1$. To demonstrate this we used a series of 10 000 random points generated from the Pareto distribution $F(y) = y^\kappa$, $0 \le y \le 1$ with shape parameter $\kappa = 1/8$. Here we expect that $D_2(m) = 0.25\, m$. In Figure 3 we have plotted the estimated correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$

(lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8. It should be noted that the length scales ε in this figure, as well as in all subsequent figures, are normalized (by rescaling data values in the interval [0, 1]). The empirical results in Figure 3, agree perfectly with the theoretical expectations ($D_2(1) = 0.25$, $D_2(2) = 0.5$, etc.).

Non-symmetric J-shaped distribution functions with large positive coefficients of skewness are the most common in hydrologic processes on fine time scales (e.g., hourly or daily), which are the most important scales when investigating the presence of determinism. Therefore, the correlation dimensions estimated from hydrologic data series do not correspond to the actual topological dimensions of the ‘attractors’.

3.3 Effect of intermittency

Things are even worse when examining rainfall series, which on fine and intermediate time scales (e.g., finer than monthly) are characterized by the presence of zeros. As shown in Appendix 1, when the probability of having zero values is finite, the correlation dimension $D_2(m)$ for any m is exactly zero. This is demonstrated in Figure 4 where we have plotted the correlation sums from a series of 10 000 independent random values 80% of which are generated from the uniform distribution and the remaining 20% are zeros, located at random. Clearly the slopes of the correlation sums are zero for small length scale ε for all embedding dimensions, except for the very large ones (7 and 8) where the zero slope is not emerging due to insufficient number of points in the data set.

Therefore, looking for correlation dimensions in a fine scale rainfall series is totally useless: the correlation dimension is simply zero for any embedding dimension. Positive estimated dimensions, such as those in the range 0.95-2.5 (items #6 and #12 in Table 1) simply indicate that a wrong range of the scale length ε was used. For example, had we estimated the correlation dimension in Figure 4 around $\varepsilon = 10^{-2}$, the resulting D_2 would be in the range 0.2 to 1.5 for embedding dimensions 1 to 5. Note that, by definition (equations (11) and (12)) the correlation dimension is theoretically determined for $\varepsilon \to 0$, which means that in practice the lowest possible region of the length scale must be used in estimations.

The problems of intermittency are not exhausted to rainfall series that contain zeros. Streamflow series display a kind of intermittency, too, as the flow shifts among different regimes, low and regular flows, and floods. For such kinds of data series, that exhibit intermittency without including zeros, *Graf von Hardenberg et al.* [1997b] have shown that the standard algorithms fail to estimate correctly the dimensions of processes characterized by intermittency, while giving no warning of their failure. In addition, they demonstrated that the Grassberger-Procaccia algorithm, applied on a time series from a composite chaotic system with randomly driven intermittency, estimates a very small dimension (e.g. $D_2 = 1$ or smaller) although the actual dimension of the system is infinite due to the random character of intermittency. Finally, they proposed ways to refine the algorithm so as to obtain correct results. The simplest of them is to filter the data by excluding all the delay vectors $\mathbf{x}$ having at least one component $x_i < c$, where c an appropriate cutoff value (typically a small percentage, e.g. 5%, of the average of the data series) that leaves out all "off" data points of the intermittent time series. This simple algorithm was proven very effective. It must be noted, however, that it reduces dramatically the number of data points, especially for large embedding dimensions, and it is well-known that the number of data points is a crucial issue in estimating dimensions, as it will be further discussed just below, in section 3.4.

The results of *Graf von Hardenberg et al.* [1997a, b] have not been given attention in hydrologic applications, although hydrologic processes of central interest such as rainfall and runoff are intermittent. This may be a source of significant errors in hydrologic applications such as those summarized in Table 1, which as we demonstrate in next sections acts synergistically with other sources of errors, thus resulting in totally flawed results.

The effect of intermittency is closely related to the effect of an asymmetric distribution function. A J-shaped distribution that is defined for positive values of the variable and has a high coefficient of skewness produces random points whose largest percentage are close zero whereas a small number of points can take extremely large values. This can be interpreted as virtually equivalent to intermittence. Therefore, the methods proposed by *Graf von Hardenberg et al.* [1997b] to recover from flawed values of dimensions are appropriate to recover from the effect of an asymmetric distribution function, as well. This will be demonstrated in section 3.5.

3.4 Effect of data size

Kantz and Schreiber [1997, p. 242] show that we need extremely many data points to recover chaos from time series and also describe the high difficulties to identify the dynamics of high dimensional (e.g. with dimension higher than 1-2) systems. However, they avoid suggesting a specific formula to estimate the sufficient number of data points required. In hydrological applications two such formulae have been used, that by *Smith* [1988],

$$N_{\min} = 42^{m} \tag{19}$$

and an approximation of the formula due to *Nerenberg and Essex* [1990],

$$N_{\min} = 10^{2 + 0.4\, m} \tag{20}$$

The first suggests that more than 10^8 and 10^{16} data points are needed to estimate the correlation dimension for embedding dimensions m = 5 and 10, respectively. The second decreases these figures significantly to the level of 10^4 and 10^6 data points, respectively. Even in the second case, however, the required data points are too many even to allow us to think of applying the time delay embedding method for dimensions higher than 5. However, most authors, as shown in Table 1, have applied the method for embedding dimensions much higher than 5 (even up to 40) and interpreted the resulting correlation dimensions as accurate enough to assure chaotic dynamics. Generally, it is hoped that both formulae overestimate the required number of data points [e.g. *Sivakumar*, 2000]. However, no proof was ever given that the formulae overestimate the required data size.

The problem of determining the data size is not in fact too difficult, as it can be reduced to a typical statistical problem. When we attempt to show that a time series originates from a low-dimensional deterministic system rather than a stochastic system, it is natural to make the null hypothesis that it originates from a stochastic system and then to reject this hypothesis. Under this null hypothesis, we can assume that the correlation sum for any length scale ε and any embedding dimension m is

$$C_2(\varepsilon, m) = [C_2(\varepsilon, 1)]^m. \tag{21}$$

In a stochastic system, $C_2(\varepsilon, m)$ expresses a probability: the probability that the distance of two points is less than ε. Using classic statistical techniques we show in Appendix 2 that the required data size to estimate $C_2(\varepsilon, m)$ is

$$N_{\min} = \sqrt{2}\,(z_{(1+\gamma)/2} / c)\,[C_2(\overline{\varepsilon}, 1)]^{-m/2} \tag{22}$$

where z_a is the a-quantile of the standard normal distribution, γ is a confidence coefficient, c is the acceptable relative error in the estimation of $C_2(\varepsilon, m)$ and $\overline{\varepsilon}$ is the highest possible length scale that suffices to accurately estimate the correlation dimension for embedding dimension 1 (meaning that for $\varepsilon > \overline{\varepsilon}$ becomes inaccurate). We can observe that the proposed formula (22) becomes identical to (20) if we assume (as typically in statistics) a confidence coefficient γ = 0.95 for which $z_{(1+\gamma)/2} = 1.96$, an acceptable error c = 3% and a sufficient $C_2(\overline{\varepsilon}, 1) = 0.15$ (indeed, $2^{0.5}\,(1.96 / 0.03)\,0.15^{-m/2} = 10^{1.97 + 0.41m} \approx 10^{2 + 0.4\,m}$). However, (22) is a more general equation and the appropriate values of c and $C_2(\overline{\varepsilon}, 1)$ need to be more carefully selected.

We will demonstrate this result and its application using an example with a totally random system. Specifically, we use a sequence of 10 000 random numbers from the Weibull distribution with shape parameter κ = 1/8 (and scale parameter 1). We know from the discussion of section 3.2 that, although the system is random, the correlation dimension $D_2(m)$ does not equal the embedding dimension m, but rather is $2\ \kappa\ m = m\ /\ 4$. In addition, since we know the probability distribution function, it is easy to calculate numerically (using equations (14) and (10)) the correlation sum $C_2(\varepsilon, 1)$ and the local slope $d_2(\varepsilon, 1)$ for any length scale ε. Then from (21) we can calculate $C_2(\varepsilon, m)$ and $d_2(\varepsilon, m)$ for any embedding dimension m. These theoretical $C_2(\varepsilon, m)$ and $d_2(\varepsilon, m)$ have been plotted in Figure 5 as continues curves. We observe from the lower panel of Figure 5 that, indeed, $D_2(1) = 0.25$, but the curve $d_2(\varepsilon, 1)$ raises very slowly from $d_2(1, 1) = 0$ to its limit value $d_2(0, 1) = D_2(1) = 0.25$, so that even for ε as low as 10^{-10} the theoretical value $d_2(10^{-10}, 1) = 0.18$ i.e., 28% smaller than the correlation dimension. Only at $\varepsilon = 10^{-20}$ becomes 0.245 (only 2% smaller than the correlation dimension). Thus, we may assume that the highest acceptable ε is $\overline{\varepsilon} = 10^{-20}$ and from the upper panel of Figure 5 we conclude that $C_2(\overline{\varepsilon}, 1) = 0.0011$.

Until now we did not use the generated time series at all, nor any statistical approach. Now, let us make the statistical calculations. We assume an acceptable statistical error c in the estimation of $C_2(\varepsilon, m)$ equal to 1%. This is safe enough and is not too small as may seem at first glance: as demonstrated in Appendix 2, it corresponds to a much larger statistical error in $d_2(\varepsilon, m)$, which may be as high as 20%; this must be considered in addition to the "theoretical" error 2% discussed in the previous paragraph. Thus, the required number of points is $N_{\min} = 2^{0.5} \times (1.96\ /\ 0.01) \times 0.0011^{-m\,/\,2} = 10^{2.44\,+\,1.48m} = 30^{1.65\,+\,m}$. This results in values of $N_{\min}$ much higher than those obtained from (20) and closer to those obtained by (19). (In fact the results of the current analysis are higher than those of (19) unless $m > 17$.) For instance, for $m = 1, 2, 5$ and 10 we get $N_{\min} = 8\ 350$, $252\ 000$, 6.9×10^9, and 1.7×10^{17}, respectively. This obviously means that is totally impractical to estimate correlation dimensions even for small dimensions, not only because of the difficulty to get such a large sample size (in our example this is not so important because data is synthesized) but also because of the huge amount of calculations required (note that the number of comparisons is in fact proportional to $N_{\min}{}^{2m}$).

Because the actual data size in our example $N = 10\,000$ is greater than $N_{\min} = 8\,350$ for $m = 1$, we can get a reliable estimate of $C_2(\varepsilon, 1)$ and $d_2(\varepsilon, 1)$ for ε even smaller than $\overline{\varepsilon} = 10^{-20}$ down to a critical value $\underline{\varepsilon}_1$. This can be estimated from (20) by replacing $N_{\min}$ with N and $\overline{\varepsilon}$ with $\underline{\varepsilon}_1$. Solving then for $C_2(\overline{\varepsilon}, 1)$ we find for $m = 1$, $C_2(\underline{\varepsilon}_1, 1) = 2\,[z_{(1+\gamma)/2} / (c\,N)]^2$. In our example, $C_2(\underline{\varepsilon}_1, 1) = 0.000768$, which, according to the graph of the upper panel of Figure 5 (after a small extrapolation) corresponds to $\underline{\varepsilon}_1 = 2.1\times10^{-21}$.

If the same data size N was used for all embedding dimensions, as is the case in most applications including this example, then the same critical value of C_2 applies to all embedding dimensions, i.e.,

$$C_2(\underline{\varepsilon}_m, m) = C_2(\underline{\varepsilon}_1, 1) = 2\,[z_{(1+\gamma)/2} / (c\,N)]^2 \tag{23}$$

This has been plotted as a dashed straight line in the upper panel of Figure 5. This line is critical for our estimations as all points of $C_2(\varepsilon, m)$ lying below this line do not have the required accuracy. The intersections of this line with the different curves $C_2(\varepsilon, m)$ determine the critical $\underline{\varepsilon}_m$ for each embedding dimension m. Given $\underline{\varepsilon}_m$ we can find the corresponding $d_2(\underline{\varepsilon}_m, m)$ so we can plot a critical curve in the lower panel of Figure 5 (dashed line), above which all points do not have the required accuracy. We must note that this example was structured based on the known probability distribution function of the variable. However the method developed can be applied even when the distribution function is not known, as we will see in next examples.

In conclusion, the proposed approach to determine the required data size or, equivalently, the adequacy of estimations for a given data size, involves two characteristic length scales: the upper limit $\overline{\varepsilon}$, which is common for all embedding dimensions, and the lower limit $\underline{\varepsilon}_m$ which is an increasing function of dimension. The required data size $N_{\min}$ for embedding dimension

m is determined setting $\underline{\varepsilon}_m = \overline{\varepsilon}$, whereas for a given N an estimation is accurate when $\underline{\varepsilon}_m \le \overline{\varepsilon}$.

Furthermore, the limits $\underline{\varepsilon}_m$ and $\overline{\varepsilon}$ can be determined in a geometrical manner even without using the data size N. The steps are the following.

1. Make plots of $C_2(\varepsilon, m)$ and $d_2(\varepsilon, m)$ for several embedding dimensions m.
2. In the plot of $d_2(\varepsilon, 1)$ (i.e., for embedding dimension 1) locate a region where $d_2(\varepsilon, 1)$ becomes constant and relatively smooth. Set $\overline{\varepsilon}$ and $\underline{\varepsilon}_1$ the upper and lower limit of this area, respectively. (Above $\overline{\varepsilon}$, $d_2(\varepsilon, 1)$ is not constant and below $\underline{\varepsilon}_1$ it becomes too rough).
3. From the plot of $C_2(\varepsilon, 1)$ determine $C_2(\underline{\varepsilon}_1, 1)$.
4. Set $C_2(\underline{\varepsilon}_m, m) = C_2(\underline{\varepsilon}_1, 1)$ and determine $\underline{\varepsilon}_m$ for each m.
5. For those m where $\underline{\varepsilon}_m \le \overline{\varepsilon}$ and $d_2(\varepsilon, m)$ is relatively constant in the interval $(\underline{\varepsilon}_m, \overline{\varepsilon})$, determine $D_2(m)$ as the average $d_2(\varepsilon, m)$ on this interval. For those m where $\underline{\varepsilon}_m > \overline{\varepsilon}$, $D_2(m)$ cannot be determined.

If for any reason the data size is different for different embedding dimensions (e.g. N_m), the equation in step 4 should be replaced by

$$C_2(\underline{\varepsilon}_m, m) = C_2(\underline{\varepsilon}_1, 1)\,(N_1 / N_m)^2. \tag{24}$$

A geometrical view of the procedure is possible by plotting the curves $\varepsilon = \overline{\varepsilon}$ and $\varepsilon = \underline{\varepsilon}_m$ in both diagrams of $C_2(\varepsilon, m)$ and $d_2(\varepsilon, m)$. This will become clearer in next examples. In the example of Figure 5 it is clear that only $D_2(1)$ can be estimated with $N = 10\,000$ points, provided that $\overline{\varepsilon} = 10^{-20}$. Had we, for instance, accepted a larger $\overline{\varepsilon} = 10^{-10}$ we would able to estimate $D_2(2)$, $D_2(3)$ and $D_2(4)$ as well, as becomes apparent by observing the dashed curve in the lower panel of Figure 5. However, the cost we would have to pay in this case would be the underestimation of dimensions by 28%, as discussed above, which notably is due to theoretical rather than statistical reasons.

3.5 Effect of autocorrelation

Hydrologic time series, especially on fine time scales, are characterized by high autocorrelation coefficients. Autocorrelation in stochastic processes may be misleadingly interpreted as low dimensional determinism when applying the standard algorithms for estimating dimensions. Examples of a highly autocorrelated stochastic processes (including fractional Gaussian noise and other simpler linear and nonlinear processes) in which the naïve application of the standard methods leads erroneously to low dimensional attractors (down to 1), have been offered by *Osborne and Provenzale* [1989]; *Theiler* [1991] and Provenzale et al. [1992] (see also *Tsonis* [1992, p. 174]).

The choice of a larger number of data points may not suffice to avoid such misleading results. Another important issue is the appropriate selection of the time delay τ in constructing delay vectors. Several authors have discussed this [see among others *Tsonis*, 1992, pp, 151-156; *Abarbenel et al.*, 1993; *Kantz and Schreiber*, 1997, pp. 130-134; *Sivakumar*, 2000]. The most common approach is to choose as τ the time where the autocorrelation function decays to 1 / e, whereas e in the base of the natural logarithms. Other options are to choose the time where the first minimum of the time delayed mutual information is located, or to optimize it inside the interval defined by the times of the 1 / e decay of autocorrelation and the minimum of mutual information. As an additional means of alleviating the effect of temporal correlation is to exclude delay vectors that are close in time. This is attained by adopting a relatively high value of w in equation (15) that is used for the estimation of correlation sums.

The effect of autocorrelation may act synergistically with the effect of an asymmetric distribution function and the effect of data size. To demonstrate this we considered a data series of 10 000 autocorrelated values with J-shaped distribution function. This was generated in the following manner: For the data point y_n, 8 random numbers were generated at a first step from the Pareto distribution with shape parameter 1/8 and at a second step the random number whose logarithm was nearest to ln y_{n-1} was chosen as y_n. This technique resulted in a series with a Markovian autocorrelation structure with lag one autocorrelation 0.72 and approximately Pareto distribution with shape parameter $\kappa = 0.44$. Therefore we expect that the correlation dimension in m dimensions of this series will be $D_2(m) = 2\ \kappa\ m = 0.88\ m$. The empirical estimates of the correlation sums and their local slopes are shown in Figure 6. These estimates were based on delay time $\tau = 4$, which corresponds to the 1 / e (= 0.37) decay of the autocorrelation function. We observe that the empirical correlation dimension for $m = 1$ agrees perfectly with the theoretical expectation be $D_2(1) = 0.88$. However the empirical $D_2(2)$ is around 1, significantly less than the expectation 1.76. The technique proposed in section 3.4 for assessing the accuracy of empirical estimation suggests that we cannot have accurate estimations of correlation dimensions for $m > 2$, as demonstrated graphically in Figure 6. If we ignored this and considered all estimated dimensions as accurate, we would conclude that correlation dimensions, estimated for ε in the interval $(10^{-4}, 10^{-3})$, saturate at about 1. Thus, we would claim that a purely stochastic system is a low-dimensional deterministic system.

The recover from this inaccurate result we can try a higher τ. However, to show the synergistic action of the several effects we chose another recovery technique. Specifically, we focused on the effect of the high skewness which can be remedied using the method due to *Graf von Hardenberg et al.* [1997b] (discussed in section 3.3) of cutting off the very small values. Applying a cutoff threshold 0.01, we determined the correlation sums and local slopes shown in Figure 7 (upper and lower panel, respectively). Clearly here, we observe that for $m = 1$ and 2, $D_2(m) = m$, whereas for higher dimensions, although accurate estimations are not possible, the figures indicate a tendency for high dimensions. Thus, the cutoff technique helps to avoid erroneous results in this example.

4. Real world examples

In light of the analyses of sections 3, in this section 4 we examine some real world hydrometeorological series, which include rainfall on daily, fine sub-daily, and monthly scale (sections 4.1, 4.2, and 4.3, respectively) relative humidity (section 4.4) and streamflow (section 4.5).

4.1 Daily rainfall series

As explained in section 3.1, the role of rainfall is crucial in investigating chaos in hydrological processes, since rainfall is the input that mobilizes all other hydrologic processes in a catchment. In section 3.1 we also presented some arguments that the rainfall process cannot be low-dimensional deterministic without applying any algorithm to determine dimensions. However, it may have some interest to apply the standard algorithm to some historical rainfall data series. Several such series were examined and the results are in all cases similar. Here we present the results for one series, the daily data at the Vakari raingage, western Greece. This raingage is located in one of the wetter parts of Greece with 40% rainy days and a mean annual rain depth approaching 1700 mm. (According to the arguments of section 3.1 the wetter the climate regime the greater is the hope of lower dimensionality of the attractor). More than 31 years or 11 476 daily data values were available. Among these years, the maximum dry spell length is 47 days, that is, 2.5 times smaller that that of the Athens raingauge discussed in section 3.1. The mean, standard deviation and coefficient of skewness of the data record are 4.59 mm, 11.90 mm and 4.59, respectively. Had the zero values excluded from the record, these statistics would be 11.38, 16.55, and 2.96, respectively. In any case, the skewness is very high and the distribution is J-shaped. The lag-one autocorrelation of the series is 0.35, which means that a delay $\tau = 1$ would suffice. Plots of the delay representation of the series in two and three dimensions are shown in Figure 2. However, for the application of the Grassberger-Procaccia algorithm we chose a much higher (and thus safer) value, $\tau = 12$, which we located as the position where the autocorrelation function has its first minimum.

In Figure 8 we have plotted the correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from the this daily rainfall series. As expected (and already discussed in section 3.1) due to the presence of zeroes in the data series the local slopes for all embedding dimensions become zero for small length scales ($\varepsilon \leq 0.0004$). Thus, this figure says nothing about the capacity dimension of the 'attractor' of the rainfall process. If we incorrectly ignored the small length scales and instead, chose length scales in the region 0.01-0.1, we would come up with small positive dimensions, not exceeding 1.5 even for embedding dimensions 8. If we also continued the plots for embedding dimensions 10, 20, 30 and so on, totally ignoring the astronomical number of data points required to do estimations in these dimensions (as is the case in some of the articles of Table 1) it is very probable that we would conclude that there is a low dimensional chaotic attractor here with dimension 1.5. This, however, would be a totally erroneous result.

It is interesting to see what happens with this data series if we exclude zero data values and apply the algorithm due to *Graf von Hardenberg et al.* [1997b]. This is shown in Figure 9, where again we observe that the local slopes $d_2(\varepsilon, m)$ become zero for small scale lengths. In this case, this is the result of round-off errors in the data values, rather than a theoretically consistent result. Specifically, 5% of the values have been rounded as 0.1 mm (which is the limit of the measuring device), 4% as 0.2 mm, 3% as 0.3 mm and so on. This is equivalent to having finite probabilities of occurrence of these values, which in turn, as discussed in section 3.1, results in zero slope of the correlation sum. The main difference of Figure 9 from Figure 8 is that the even for large scale lengths the local slopes tend to more reasonable values, i.e., to about 0.7 for $m = 1$, 1.4 for $m = 2$, etc. To minimize the effect of round-off errors we also performed another application of the algorithm by *Graf von Hardenberg et al.* [1997b]

excluding data values that are less than 2 mm. In this case the data size becomes too low to allow for any accurate estimation but clearly shows that the correlation dimension $D_2(m)$ tends to the embedding dimension m, which means that the time series has a stochastic character.

4.2 Storm data

If the presence of zeros in a rainfall time series is an strong obstacle to analyze the presence chaos, one may think that going to a much finer scale and limiting the analysis strictly on a rainy period (a single storm) he or she could identify the deterministic chaos in there. The idea of a deterministic evolution of a storm has been favored long before hydrologists be involved with chaos. For example, *Eagleson* [1970, p. 184] states "The spacing and sizing of individual events in the sequence is probabilistic, while the internal structure of a given storm may be largely deterministic".

To explore this idea we used a storm time series measured with high temporal resolution. This data set corresponds to one of several storms that were measured by the Hydrometeorology Laboratory at the University of Iowa using devices that are capable of high sampling rates [*Georgakakos et al.*, 1994]. The data is available on the Internet from ftp.iihr.uiowa.edu. Specifically, the data set used is that of the event labeled Rain 1, which occurred during 2-3 December 1990. This data set was the subject of several advanced and extensive analyses including multifractal analysis and multiplicative cascades [*Carsteanu and Foufoula-Georgiou*, 1996] and wavelet analysis [*Kumar and Foufoula-Georgiou*, 1997].

The duration of this storm was almost 27 h and the rain depth was measured every 10 s, so that the data set contains 9679 data points. The total depth is 104.9 mm, and the mean, maximum and minimum 10 s intensity are 3.89, 118.74, and 0.07 mm/h, respectively. The standard deviation of the 10-second intensities is 6.16 mm/h (1.58 times the mean) and their coefficient of skewness is 4.83. The distribution function is J-shaped and the gamma distribution function with a shape parameter 0.40 gives an acceptable fit to the data series. The autocorrelation is very high. For lags 1, 100 and 500 is 0.88, 0.48 and 0.15, respectively, and only at lag 850 becomes zero. For high frequencies (> 4×10^{-3} cycles per second) the power spectrum is approximately a power function of frequency with an exponent 1.63 (estimated by *Georgakakos et al.* [1994]).

The correlation sums $C_2(\varepsilon, m)$ of this time series for $\tau = 500$ and their local slopes $d_2(\varepsilon, m)$ are plotted in Figure 10 versus length scale ε for embedding dimensions $m = 1$ to 8. Again here we observe zero slopes for low scale lengths. These again are due to round-off errors that artificially result in equal values: for example 217 values are 0.09 mm/h and 169 are 0.08 mm/h. If we ignore the regions with zero slopes, and apply the statistical reasoning exhibited in section 3.4, we find that for the plot of $m = 1$ the upper limit for adequate estimations is $\overline{\varepsilon} = 0.008$ and the lower limit for accurate estimations is $\underline{\varepsilon}_1 = 0.0014$. For $m = 2$, we find from (24) that $\underline{\varepsilon}_2 = 0.0072 < \overline{\varepsilon}$, whereas for all greater m, $\underline{\varepsilon}_m > \overline{\varepsilon}$. Thus, $D_2(m)$ can be estimated only for $m = 1$ and 2, and the estimated values are $D_2(1) = 0.69$ and $D_2(2) = 1.00$. Given that the shape parameter of the gamma distribution is 0.40, the expected values for an entirely random series are 0.80 and 1.60 for $m = 1$ and 2 respectively. In any case, these results do not support nor prohibit the existence of low-dimensional deterministic dynamics.

As an additional analysis, we applied the *Graf von Hardenberg et al.* [1997b] algorithm excluding data values smaller than 1% of the maximum value and plotted the resulting correlation sums and local slopes in Figure 11. Now it can be observed that $D_2(1) = 1$ and for higher dimensions, although no accurate estimations can be obtained, it is apparent the tendency that $D_2(m) = m$, which indicates a stochastic behavior.

4.3 Monthly rainfall series

It has been found that in some cases with many degrees of freedom, only a few of them remain active due to some collective behavior [*Kantz and Schreiber*, 1997, p. 34].

Specifically, many systems are composed of a huge number of internal microscopic degrees of freedom, but nevertheless produce signals which are found to be low dimensional. The coupling between the different degrees of freedom and an external field of some kind, lead to collective behavior which is low dimensional. The reason is that most degrees of freedom are either not excited at all or "slaved". [*Kantz and Schreiber*, 1997, p. 239].

By analogy, it may have meaning to study a hydrologic process on a coarse time scale and try to identify chaos. Even if the system on a fine scale appears as random, one may think of some collective behavior on the coarser scale, which could result in a low-dimensional attractor.

Here we present the results for one series on a coarse timescale, the monthly rainfall in at the station of the National Observatory in Athens, which is the longest rainfall record in Greece [see *Koutsoyiannis and Baloutsos*, 2000]. This corresponds to a dry climate with about 400 mm annual rainfall; in 9% of the months the rainfall is zero. More than 132 years or 1586 monthly data values were available (August 1859 to September 1991). The mean, standard deviation and coefficient of skewness of the data record are 32.9 mm, 36.0 mm and 1.75, respectively. Had the zero values excluded from the record, these statistics would be 36.4, 36.2, and 1.70, respectively. Despite of the large skewness, the distribution is bell-shaped. The autocorrelation coefficient of the series is 0.32 for lag one and decays quickly, so that it becomes negative for lag three.

The correlation sums and their local slopes of this series, excluding zero points and using delay $\tau = 1$, are plotted in Figure 12 versus length scale ε for embedding dimensions $m = 1$ to 8. Due to the small record size only the estimate of $D_2(1)$ is accurate (as verified from the graphical application of the procedure described in section 3.4 shown in Figure 12) and is about 1. For higher dimensions no accurate estimations can be obtained, but again the tendency is that $D_2(m) = m$, which indicates a stochastic behavior.

4.4 Relative humidity series

Since we found difficulties in identifying chaos in rainfall on all timescales, it could be a good idea to move to another related process towards the direction of meteorology. The meteorological variable most closely related to rainfall is the relative humidity since when it rains, it approaches saturation (i.e., the value 100%). The data series we used is the relative humidity of the period 1 December 1998 to 4 February 2001 on hourly scale (18 888 data values) and comes from the meteorological station of the National Technical University in Athens (made available on the Internet at www.hydro.ntua.gr/meteo/); a few missing values were filled in by linear interpolation in time. Obviously, the relative humidity series is totally free from zeros and intermittency, which makes its study easier. The mean, standard deviation and coefficient of skewness of the data record are 60.2%, 17.2% and –0.26, respectively, whereas the minimum and maximum values are 12.3% and 99.0%. The distribution is bell-shaped. The autocorrelation coefficient of the series is as high as 0.97 for lag one and decays slowly, so that it becomes smaller than 1/e only for lag 108.

The correlation sums $C_2(\varepsilon, m)$ of this time series for $\tau = 108$ and their local slopes $d_2(\varepsilon, m)$ are plotted in Figure 13 versus length scale ε for embedding dimensions $m = 1$ to 8. We observe on the plots of $m = 1$ that a long scaling area appears between $\overline{\varepsilon} = 0.08$ and $\underline{\varepsilon}_1 = 0.00092$. Thus, $\underline{\varepsilon}_m < \overline{\varepsilon}$, for $m \leq 4$, as shown graphically in Figure 13, which means that $D_2(m)$ can be estimated accurately for $m = 1$ to 4. The estimated values are $D_2(m) = m$, a result that again does not allow any hope for low-dimensional determinism.

4.5 Daily streamflow series

Finally, we will study the most representative hydrological process using a daily streamflow series of the Pinios River, central-eastern Greece, at the Ali Efenti gage. The data series extends through the period 3 January 1972 to 18 March 1998 (8 246 data values of which 1435 were missing data that were left unfilled). As explained in section 3.3, a streamflow series must be regarded as intermittent even if it is free from zeros. The mean, standard deviation and coefficient of skewness of the data record are 39.6 m^3/s, 56.5 m^3/s and 3.46, respectively, whereas the minimum and maximum values are 1.0 m^3/s and 553.5 m^3/s. The distribution is very asymmetric yet bell-shaped. The autocorrelation coefficient of the series is as high as 0.86 for lag one and decays slowly, so that it becomes zero only for lag 94.

The correlation sums $C_2(\varepsilon, m)$ of this time series for $\tau = 94$ and their local slopes $d_2(\varepsilon, m)$ are plotted in Figure 14 versus length scale ε for embedding dimensions $m = 1$ to 8. We observe on the plots of $m = 1$ that a scaling area appears between $\overline{\varepsilon} = 0.06$ and $\underline{\varepsilon}_1 = 0.001$, whereas for all other m, $\underline{\varepsilon}_m > \overline{\varepsilon}$, which means that an accurate estimation of $D_2(m)$ is possible only for $m = 1$; this is $D_2(1) \approx 1$. For higher embedding dimensions m, a tendency appears for $D_2(m)$ increasing with m, which again indicates a stochastic behavior.

5. Summary and conclusions

Several recent studies using methods of chaotic analysis have claimed for discovering low-dimensional determinism in hydrologic processes such as rainfall and runoff. Other studies have expressed skepticism about the results of the former ones, indicating that such results are suspect or erroneous due to naïve application of the theory and related algorithms. This paper has attempted to offer some additional insights on this debating discussion by studying several aspects of dynamical systems and their application to the characterization of the hydrologic processes. Specifically, it shows that in some cases merely the careful application of the concepts of dynamical systems, without doing any calculation, provides strong indications that hydrologic processes cannot be (low-dimensional) chaotic. Furthermore, it shows that specific peculiarities of hydrologic processes are synergistic factors that can lead to misleading conclusions regarding presence of chaos.

The arguments that are presented and studied in the paper are the following:

1. A time series that contains periods with zero values cannot be the outcome of a low-dimensional deterministic dynamical system. In this respect rainfall cannot be a (low-dimensional) chaotic process.
2. In addition, since rainfall is the input that mobilizes all other hydrologic processes in a catchment, such as streamflow, these processes cannot be chaotic, too.
3. An attractor dimension as low as 1 or even smaller, which some studies have claimed for in hydrological processes, should be directly visualized via delay representation graphs. This however, has never come into light, simply because in fact such graphs manifest space filling clouds rather than one-dimensional structures.
4. The attractor dimension must be consistent with the dimension used to embed it according to Whitney's embedding theorem. For example, if an attractor dimension was 1 or less, then a three dimensional embedding space would suffice to embed it. The fact that the required embedding dimension in some studies is as high as 10-40 simply indicates that the results are flawed.
5. The embedding theorems are in fact based on the concept of the capacity dimension whereas the standard algorithms to determine attractor dimensions use the concept of the correlation dimension. The two different dimensions are generally identical but not in hydrologic processes on fine time-scales. Specifically, it has been showed that if the distribution function is J-shaped with high skewness, as is the case with hydrologic processes on fine time-scales, the correlation dimension is smaller than the capacity dimension. This may lead to misleadingly small estimated dimensions.
6. Intermittency (which is apparent in hydrologic processes – not only in rainfall but in streamflow as well) is another factor that can result in a misleading low attractor dimension even in stochastic, i.e. infinite dimensional, systems. This result has been known from earlier works but has not been given the required attention in hydrological studies investigating chaos.

7. Another known issue that has not been given the required attention in similar studies is the fact that we need extremely many data points to recover chaos from time series, which are hardly available for hydrologic processes. Using statistical reasoning we have shown that the required data size in hydrologic time series may be even more exceptionally high due to the asymmetric distribution functions. We also propose a methodology for estimating the required number of data points for a certain embedding dimension or, conversely, the maximum allowed embedding dimension to accurately recover chaotic dynamics for a given number of data points. It turns out that for typical applications, accurate estimations can be obtained only for embedding dimensions 1-2, which are hardly sufficient to recover dynamics.
8. The high autocorrelation that characterizes many hydrologic processes, mostly on fine time-scales, is another factor that, acting synergistically with other factors such as asymmetric distribution and insufficient data size, may be misleadingly interpreted as low dimensional determinism when applying the standard algorithms for estimating dimensions.

All these arguments have been demonstrated using appropriately synthesized theoretical examples. Finally, in light of the theoretical analyses and arguments, typical real-world hydrometeorological time series, which include rainfall on daily, fine sub-daily, and monthly scale, relative humidity, and streamflow, have been explored and none of them is found to indicate the presence of chaos but, rather, all must be regarded as the outcomes of stochastic systems.

Appendix 1: Theoretical investigation of correlation dimension of asymmetric processes

Let Y_n be a random process on discrete time n with all Y_n (n = 1, 2, …) independent identically distributed positive variables with distribution function $F(y)$ and density $f(y)$. We assume that $y > 0$ (as happens with all hydrologic variables) and also $y < \xi$ where the upper bound ξ could be finite or infinite. We will study the correlation dimension for embedding dimension $m = 1$. We consider a partition of the y domain with length scale ε. Applying (8) and (11) and observing that $p_i = F(i\,\varepsilon) - F((i-1)\varepsilon)$ we find

$$D_2 = \lim_{\varepsilon \to 0} \left\{ \frac{1}{\ln \varepsilon} \ln \sum_{i=1}^{v(\varepsilon,\, \xi)} [F(i\,\varepsilon) - F((i-1)\varepsilon)]^2 \right\} \tag{25}$$

where $v(\varepsilon, \xi)$ is the smallest integer that is greater than or equal to ξ / ε. For small values of ε $F(i\,\varepsilon) - F((i-1)\varepsilon) = f(i\,\varepsilon)\,\varepsilon$ and therefore when $\varepsilon \to 0$,

$$D_2 = 1 + \lim_{\varepsilon \to 0} \left[\frac{1}{\ln \varepsilon} \ln \sum_{i=1}^{v(\varepsilon,\, \xi)} f^2(i\,\varepsilon)\,\varepsilon \right] \tag{26}$$

Also, when $\varepsilon \to 0$,

$$\sum_{i=1}^{v(\varepsilon,\, \xi)} f^2(i\,\varepsilon)\,\varepsilon \to \int_0^{\xi} f^2(y)\, dy =: B \tag{27}$$

Obviously, if B converges then $D_2 = 1$. This case is the most commonly met, since in most cases $f(y)$ has a finite value for the entire domain of y. If $f(y)$ is finite, B converges even if ξ is infinite. This is understood if we observe that in this case, there exists a $y_0 > 0$ so that $f(y) < 1$ for any $y > y_0$ and therefore $f^2(y) < f(y)$. Thus, the integral $f^2(y)$ in $[x_0, \infty)$ is finite, and since $f(y)$ is finite everywhere, the integral in $[0, \infty)$ will be finite, too. Consequently the limit in (26) becomes zero and $D_2 = 1$.

Now we consider the case that B does not converge. There are two possible necessary conditions that may lead to this case: $f(y)$ tends to ∞ either when y tends to 0 or y tends to ξ

where ξ is finite. We concentrate on the first condition, which is the most interesting as far as hydrological processes are concerned. In this case only the first term ($i = 1$) of the sum in (25) is significant so that

$$D_2 = 2 \lim_{\varepsilon \to 0} \frac{\ln [F(\varepsilon)]}{\ln \varepsilon} \tag{28}$$

Applying de l'Hospital's rule twice we get

$$D_2 = 2 \lim_{\varepsilon \to 0} \frac{\varepsilon f(\varepsilon)}{F(\varepsilon)} = 2 + 2 \lim_{\varepsilon \to 0} \frac{\varepsilon f'(\varepsilon)}{f(\varepsilon)} \tag{29}$$

where $f'(\ )$ is the derivative of $f(\ )$.

Now let us view a few examples. First we consider the Pareto distribution, in which

$$F(y) = (y / a)^{\kappa},\ f(y) = (\kappa / a)\,(y / a)^{\kappa - 1},\ f'(y) = (\kappa - 1)(\kappa / a^2)\,(y / a)^{\kappa - 2},\ \ 0 \le y \le a \tag{30}$$

Here $\xi = a$. The integral B converges to $(\kappa^2 / a) / (2\,\kappa - 1)$ when $\kappa > 1 / 2$ and diverges when $\kappa < 1 / 2$. Therefore, for $\kappa > 1 / 2$, $D_2 = 1$, whereas for $\kappa < 1 / 2$,

$$D_2 = 2 + 2 \lim_{\varepsilon \to 0} \frac{(\kappa - 1)(\kappa / a^{\kappa})\,\varepsilon^{\kappa - 1}}{(\kappa / a^{\kappa})\,\varepsilon^{\kappa - 1}} = 2\,\kappa \tag{31}$$

We note that the coefficient of skewness of this distribution is

$$C_s(\kappa) = \frac{2\,(1 - \kappa)\sqrt{2 + \kappa}}{(3 + \kappa)\sqrt{\kappa}} \tag{32}$$

which means that the correlation dimension is smaller than 1 when the coefficient of skewness is greater than $C_s(1 / 2) = 0.639$.

In our second example we consider the gamma distribution, in which

$$f(y) = [1 / a\,\Gamma(\kappa)]\,(y / a)^{\kappa - 1}\,\mathrm{e}^{-y / a},\ f'(y) = [1 / a^2\,\Gamma(\kappa)]\,(\kappa - 1 - y / a)\,(y / a)^{\kappa - 2}\,\mathrm{e}^{-y / a},\ \ y > 0 \tag{33}$$

The integral B converges to $\Gamma(\kappa - 1/2) / [2 \sqrt{\pi}\, a) \Gamma(\kappa)]$ when $\kappa > 1 / 2$ and diverges when $\kappa < 1 / 2$. Therefore, for $\kappa > 1 / 2$, $D_2 = 1$, whereas for $\kappa < 1 / 2$,

$$D_2 = 2 + 2 \lim_{\varepsilon \to 0} \frac{[1 / a^2\, \Gamma(\kappa)]\, \varepsilon\, (\kappa - 1 - \varepsilon / a)\, (\varepsilon / a)^{\kappa - 2}\, e^{-\varepsilon / a}}{[1 / a\, \Gamma(\kappa)]\, (\varepsilon / a)^{\kappa - 1}\, e^{-\varepsilon / a}} = 2\, \kappa \tag{34}$$

We note that the coefficient of skewness of this distribution is $C_s(\kappa) = 2 / \sqrt{\kappa}$, which means that the correlation dimension is smaller than 1 when the coefficient of skewness is greater than $C_s(1 / 2) = 2.83$.

In our third example we consider the Weibull distribution, i.e.,

$$F(y) = 1 - \exp[-(y / a)^\kappa],\ f(y) = (\kappa / a)\, (y / a)^{\kappa - 1} \exp[-(y / a)^\kappa]$$

$$f'(y) = (\kappa / a^2)\, [\kappa - 1 - (y / a)^\kappa]\, (y / a)^{\kappa - 2} \exp[-(y / a)^\kappa],\ y > 0 \tag{35}$$

The integral B converges to $(\kappa / a)\, \Gamma(2 - 1/ \kappa) / 2^{2 - 1 / \kappa}$ when $\kappa > 1 / 2$ and diverges when $\kappa < 1 / 2$. Therefore, for $\kappa > 1 / 2$, $D_2 = 1$, whereas for $\kappa < 1 / 2$,

$$D_2 = 2 + 2 \lim_{\varepsilon \to 0} \frac{(\kappa / a^2)\, \varepsilon\, [\kappa - 1 - (\varepsilon / a)^\kappa]\, (\varepsilon / a)^{\kappa - 2} \exp[-(\varepsilon / a)^\kappa]}{(\kappa / a)\, (\varepsilon / a)^{\kappa - 1} \exp[-(\varepsilon / a)^\kappa]} = 2\, \kappa \tag{36}$$

We note that the coefficient of skewness of this distribution is

$$C_s(\kappa) = \frac{2\, \Gamma^3(1 + 1 / \kappa) - 3\, \Gamma(1 + 1 / \kappa)\, \Gamma(1 + 2 / \kappa) + \Gamma(1 + 3 / \kappa)}{[\Gamma(1 + 2 / \kappa) - \Gamma^2(1 + 1 / \kappa)]^{3 / 2}} \tag{37}$$

which means that the correlation dimension is smaller than 1 when the coefficient of skewness is grater than $C_s(1 / 2) = 6.62$.

Appendix 2: Required data size to estimate attractor dimensions

It is well known that the length of the confidence interval of the estimate of a probability p from a sample with relatively high size N for a confidence coefficient γ is

$2\, z_{(1+\gamma)/2}\, [p\,(1-p)\,/\,N]^{0.5}$ where z_a is the a-quantile of the standard normal distribution. If c is the acceptable relative error in the estimation of p then

$$2\, z_{(1+\gamma)/2}\, [p\,(1-p)\,/\,N]^{0.5} = 2\, c\, p \tag{38}$$

Solving for N, we find that the minimum sample size $N_{\min}$ that is required for estimating the probability p with confidence γ and acceptable relative error c is

$$N_{\min} = (z^2_{(1+\gamma)/2}\,/\,c^2)\,(1\,/\,p - 1) \tag{39}$$

Now, if we replace p with the correlation sum (by its definition (15), $C_2(\varepsilon, m)$ is a probability) and $N_{\min}$ with $N^2_{\min}\,/\,2$ (since our sample in this case is composed of pairs of values), and also ignore 1 in the last term of (39) (assuming that p is small so that $1\,/\,p$ is much larger than 1) we find that the minimum sample size required to estimate $C_2(\varepsilon, m)$ with confidence γ and acceptable relative error c is

$$N_{\min} = (z_{(1+\gamma)/2}\,/\,c)\,[2\,/\,C_2(\varepsilon, m)]^{0.5} \tag{40}$$

For a stochastic system, combining (40) with (21) we find

$$N_{\min} = \sqrt{2}\,(z_{(1+\gamma)/2}\,/\,c)\,[C_2(\varepsilon, 1)]^{-m/2} \tag{41}$$

If we replace ε in (41) with $\overline{\varepsilon}$, the highest possible length scale that suffices to accurately estimate the correlation dimension, we get (22).

For the choice of an acceptable relative error c we must investigate the relation of the relative error in estimating $C_2(\varepsilon, m)$ with that in estimating $d_2(\varepsilon, m)$, which is our final target. We assume that the local slope is calculated from two successive values of $C_2(\varepsilon, m)$, at length scales ε_1 and $\varepsilon_2 = \alpha\, \varepsilon_1$, whose theoretical values are $C_2(\varepsilon_2, m) = \beta\, C_2(\varepsilon_1, m)$. We also assume that the empirical values depart from the theoretical ones by c each on opposite direction, i.e., $C'_2(\varepsilon_2, m) = (1 + c)\, C_2(\varepsilon_2, m)$ and $C'_2(\varepsilon_1, m) = (1 - c)\, C_2(\varepsilon_1, m)$. The theoretical local slope is

$$d_2(\varepsilon, m) = \frac{\ln[C_2(\varepsilon_2, m)] - \ln[C_2(\varepsilon_1, m)]}{\ln \varepsilon_2 - \ln \varepsilon_1} = \ln \beta \,/\, \ln \alpha \tag{42}$$

whereas the estimated slope will be

$$d'_2(\varepsilon, m) = \frac{\ln[C'_2(\varepsilon_2, m)] - \ln[C'_2(\varepsilon_1, m)]}{\ln \varepsilon_2 - \ln \varepsilon_1} \approx 2\, c \,/\, \ln \alpha + \ln \beta \,/\, \ln \alpha \tag{43}$$

where we have considered $\ln (1 \pm c) \approx \pm c$ due to the small value of c. Therefore, the relative error in $d_2(\varepsilon, m)$ is $2\, c \,/\, \ln \beta$. For $\beta = 0.9$, the relative error becomes $\approx 20\, c$ which means that a 1% error in $C_2(\varepsilon_2, m)$ can result in an error in $d_2(\varepsilon, m)$ as high as 20%.

References

Abarbanel, H. D. I., R. Brown, J. J. Sidorowich, and L. S. Tsimring, The analysis of observed chaotic data in physical systems, Rev. Mod. Phys., 65(4), 1331-1391, 1993.

Carsteanu, A., and E. Foufoula-Georgiou, Assessing dependence among weights in a multiplicative cascade model of temporal rainfall, *Journal of Geophysical Research-Atmospheres,* 101(D21), 26363-26370, 1996.

Eagleson, P. S., *Dynamic Hydrology*, McGraw-Hill, 1970.

Georgakakos, K. P., A. A. Carsteanu, P. L. Sturdevant, and J. A. Cramer, Observation and analysis of Midwestern rain rates, *J. Appl. Meteorol.*, 33, 1433-1444, 1994.

Gershenfeld, N. A., and A. S. Weigend, The future of time series: Learning and understanding, in *Time Series Prediction: Forecasting the Future and Understanding the Past*, edited by A. S. Weigend and N. A. Gershenfeld, pp. 1-70, SFI Stud. in the Sci. of Complex., Proc. Vol. XV, Addison-Wesley, Reading, Mass. 1993.

Ghilardi, P., and Rosso, R., Comment on “Chaos in rainfall”, Water Resour. Res., 26(8), 1837-1839, 1990.

Graf von Hardenberg, J., F. Paparella, N. Platt, A. Provenzale, and E. A. Spiegel, Through a glass darkly, in *Nonlinear Signal and Image Analysis*, edited by J. R. Buchler and H. Kandrup, Annals of the New York Academy of Sciences, 808, 79-98, 1997a.

Graf von Hardenberg, J., F. Paparella, N. Platt, A. Provenzale, E. A. Spiegel, and C. Tesser, Missing motor of on-off intermittency, Physical Review E, 55(1), 58-64, 1997b.

Grassberger, P., Generalized dimensions of strange attractors, *Phys. Lett.*, *97*A(6), 227-230, 1983.

Grassberger, P., and I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.*, *50*(5), 346-349, 1983.

Grassberger, P., T. Schreiber, and C. Schaffrath, Nonlinear time sequence analysis, Int. J. Bifurcation and Chaos, 1, 521, 1991.

Hübner, U., C. O. Weiss, N. Abraham, and D. Tang, Lorenz-like chaos in NH_3-FIR lasers, in *Time Series Prediction: Forecasting the Future and Understanding the Past*, edited by A. S. Weigend and N. A. Gershenfeld, SFI Studies in the Sciences of Complexity, Proc. Vol. XV, pp. 73-105, Addison-Wesley, 1993.

Jayawardena, A. W., and F. Lai, Analysis and prediction of chaos in rainfall and stream flow time series, *J. Hydrol.*, 153, 23-52, 1994.

Kantz, H., and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, 1997.

Koutsoyiannis, D., and G. Baloutsos, Analysis of a long record of annual maximum rainfall in Athens, Greece, and design rainfall inferences, *Natural Hazards,* 22(1), 31-51, 2000.

Koutsoyiannis, D., and D. Pachakis, Deterministic chaos versus stochasticity in analysis and modeling of point rainfall series, *Journal of Geophysical Research-Atmospheres,* 101(D21), 26444-26451, 1996.

Kumar, P., and E. Foufoula-Georgiou, Wavelet analysis for geophysical applications, *Reviews of Geophysics*, 35(4), 385-412, 1997.

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

Nerenberg, M. A. H., and C. Essex, Correlation dimension and systematic geometric effects, *Phys. Rev. A*, 42, 7065-7074, 1990.

Osborne, A. R., and A. Provenzale, Finite correlation dimension for stochastic systems with power-law spectra, *Physica D*. 35, 357-381, 1989.

Porporato, A., and L. Ridolfi, Nonlinear analysis of river flow time sequences, *Water Resour. Res.*, 33(6), 1353-1367, 1997.

Provenzale, A., L. A. Smith, R. Vio and G. Murante, Distinguishing between low-dimensional dynamics and randomness in measured time series, *Physica D*, 58, 31-49, 1992.

Rényi, A., *Probability Theory*, North-Holland, Amsterdam,1970.

Rodriguez-Iturbe, I., Exploring complexity in the structure of rainfall, *Adv. Water Resour.*, 14(4), 162-167, 1991.

Rodriguez-Iturbe, I., B. F. de Power, M. B. Sharifi, and K. P. Georgakakos, Chaos in Rainfall, *Water Resour. Res.*, 25(7), 1667-1675, 1989.

Rössler, O. E., An equation for hyperchaos, Phys. Lett. A, 71, 155, 1979.

Sangoyomi, T. B., U. Lall, and H. D. I. Abarbanel, Nonlinear dynamics of the Great Salt Lake: dimension estimation, *Water Resour. Res.*, 32(1), 149-159, 1996.

Sauer, T., J. Yorke, and M. Casdagli, Embedology, *J. Stat. Phys.*, 65(3/4), 579-616, 1991.

Sharifi, M. B., K. P. Georgakakos, and I. Rodriguez-Iturbe, Evidence of deterministic chaos in the pulse of storm rainfall, *J. Atmos. Sci.*, 45(7), 888-893, 1990.

Sivakumar, B., Chaos theory in hydrology: important issues and interpretations, *J. Hydrol.*, 227, 1-20, 2000.

Sivakumar, B., S.-Y. Liong, and C.-Y. Liaw, Evidence of chaotic behavior in Singapore rainfall, *J. Am. Water Resour. Assoc.*, 34(2) 301-310, 1998.

Sivakumar, B., S.-Y. Liong, C.-Y. Liaw, and K.-K. Phoon, Singapore rainfall behavior: chaotic? *J. Hydrol. Eng.*, ASCE, 4(1), 38-48, 1999.

Sivakumar, B., R. Berndtsson, J. Olsson, K Jinno, and A. Kawamura, Dynamics of monthly rainfall-runoff process at the Göta basin: A search for chaos, *Hydrology and Earth System Sciences*, 4(3), 407-417, 2000.

Sivakumar, B., R. Berndtsson, J. Olsson, and K Jinno, Evidence of chaos in the rainfall-runoff process, *Hydrological Sciences Journal*, 46(1), 131-145, 2001.

Smith, L. A., Intrinsic limits on dimension calculations, *Phys. Lett. A*, 133, 283-288, 1988.

Takens, F., Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence*, edited by D. A. Rand and L.-S. Young, lecture notes in Mathematics, 898, pp. 336-381, Spinger-Verlag, New York, 1981.

Theiler, J., Some comments on the correlation dimension of $1/f^a$ noise, Phys. Lett. A, 155, 480-492, 1991.

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

Tsonis, A. A., J. B. Elsner, and K. Georgakakos, Estimating the dimension of weather and climate attractors: Important issues on the procedure and interpretation, *J. Atmos. Sci.*, 50(15) 2249-2555, 1993.

Wang, Q., and T. Y. Gan, Biases of correlation dimension estimates of streamflow data in the Canadian prairies, *Water Resour. Res.*, 34(9), 2329–2339, 1998.

Whitney, H., Differentiable manifolds, *Ann. Math.*, 37, 645, 1936.

Wilcox, B. P., M. S. Seyfried, and T. H. Matison, Searching for Chaotic Dynamics in Snowmelt Runoff, *Water Resour. Res.*, 27(6), 1005-1010, 1991.

List of Figures

Figure 1 Delay representation of a series of 10 000 points generated from the linearly routed logistic equation (see text) in two (upper panel) and three (lower panel) dimensions.

Figure 2 Delay representation of a series of 10 000 daily rainfall depths in two (upper panel) and three (lower panel) dimensions.

Figure 3 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from a series of 10 000 independent random values with Pareto distribution with exponent 1/8.

Figure 4 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from a series of 10 000 independent random values, 80% of which are generated from the uniform distribution and the remaining are zeros (located at random).

Figure 5 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from a series of 10 000 independent random points from the Weibull distribution with shape parameter 1/8.

Figure 6 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from a series of 10 000 autocorrelated random values having approximately Pareto distribution with shape parameter 0.44.

Figure 7 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from the same series as in Figure 6 but excluding points having at least one coordinate smaller than 0.01.

Figure 8 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the daily rainfall series at the Vakari raingage.

Figure 9 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the same daily rainfall series as in Figure 8 but excluding points with zero values.

Figure 10 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the fine scale rainfall series at Iowa.

Figure 11 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the fine scale rainfall same series as in Figure 10 but excluding points having at least one coordinate smaller than 0.01.

Figure 12 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the monthly rainfall series at Athens excluding zero points.

Figure 13 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the relative humidity series at Athens.

Figure 14 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the discharge series at Ali Efenti gage at Pinios River.

Tables

Table 1 List of studies that have investigated the presence of chaos in hydrologic processes.

Item #	Reference	Data type	Location	Time scale	Data size	Time delay used	Embedding dimension	Attractor dimension
1.	*Rodriguez-Iturbe et al.* [1989]	rainfall	Genoa	weekly	7722	not mentioned	up to 8	no convergence
2.	*Rodriguez-Iturbe et al.* [1989]; *Rodriguez-Iturbe* [1991]	storm	Boston	15 s	1990	8-12	5	3.78
3.	*Sharifi et al.* [1990]	raingauge tip times	Cambridge, Massachusetts	0.01 mm	3316-4000	4-134	8-10	3.35-3.75
4.	*Wilcox et al.* [1991]	runoff	Reynolds Mountain, Idaho	daily	8800	2-16	up to 20	no convergence
5.	*Tsonis et al.* [1993]; *Tsonis* [1992, p. 168]	raingauge tip times	not reported	0.01 mm	2200	not needed	5	2.2
6.	*Jayawardena and Lai* [1994]	rainfall	3 stations in Hong Kong	daily	3650-4015	2	30-40	0.95-2.54
7.	*Jayawardena and Lai* [1994]	streamflow	2 stations in Hong Kong	daily	5840-7300	2-3	10-20	0.45-0.65

Table 1 List of studies that have investigated the presence of chaos in hydrologic processes. **(continued)**

Item #	Reference	Data type	Location	Time scale	Data size	Time delay used	Embedding dimension	Attractor dimension
8.	*Koutsoyiannis and Pachakis* [1996]	rainfall	Ortona, Florida	0.25 h to 24 h	70 000 - 2214	96-12	up to 32	no convergence
9.	*Sangoyomi et al.* [1996]	lake volume	Great Salt Lake	biweekly	3750	9	8	3.4
10.	*Porporato and Ridolfi* [1997]	streamflow	Dora Baltea (tributary of Po)	daily	14246	1	<10	<4
11.	Wang and Gan [1998]	streamflow	6 rivers in Canadian Prairies	daily	3044-30316	40-180	10	3 (interpreted to be 7-9)
12.	*Sivakumar et al.* [1998, 1999]	rainfall	6 stations in Singapore	daily	10958	7-20	12-13	1.01-1.03
13.	*Sivakumar et al.* [2000, 2001]	rainfall	Göta, Sweden	monthly	1572	3	10	6.4
14.	*Sivakumar et al.* [2000, 2001]	runoff	Göta, Sweden	monthly	1572	20	10	5.5
15.	*Sivakumar et al.* [2000, 2001]	runoff coefficient	Göta, Sweden	monthly	1572	3	13	7.8

Figures

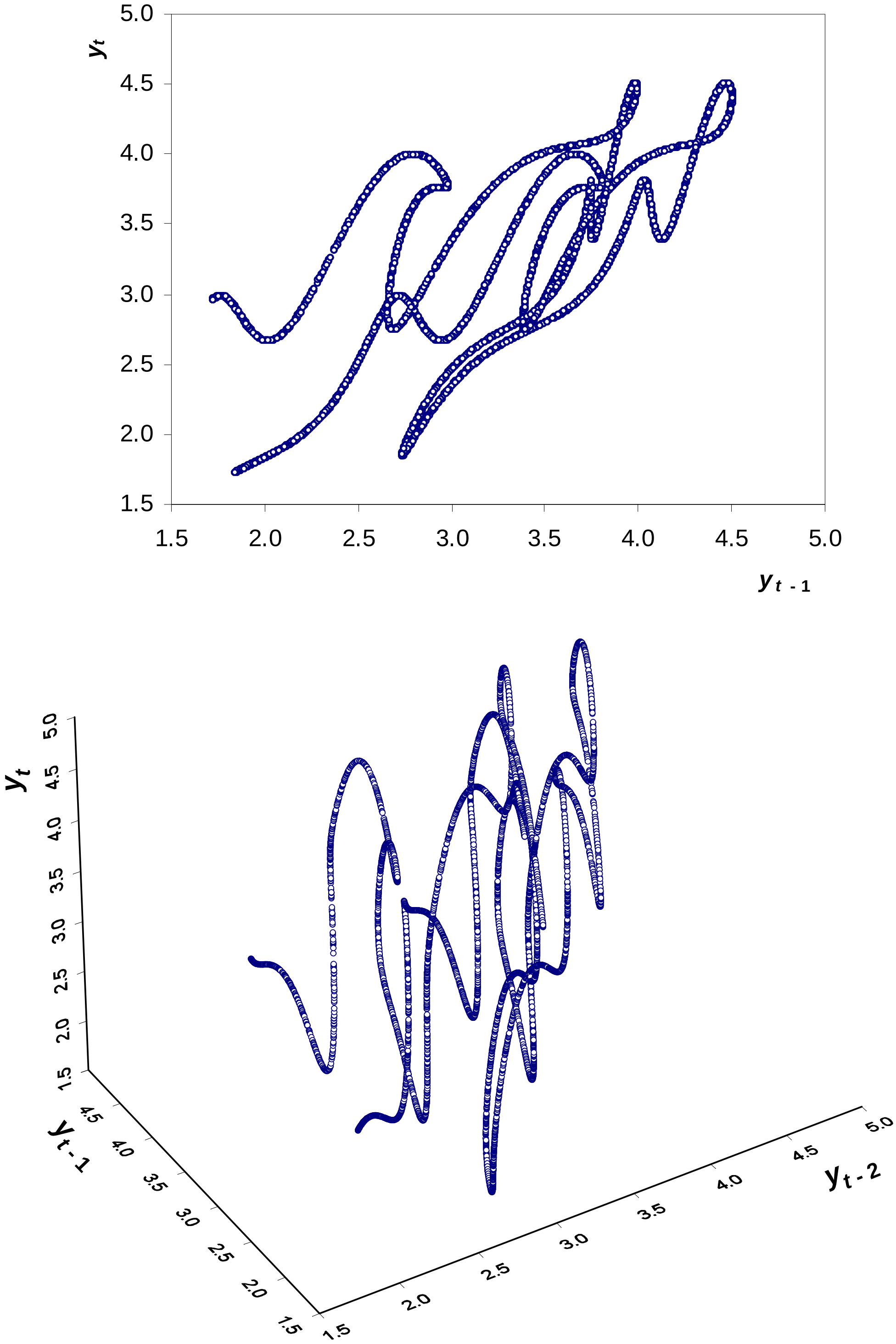

Figure 1 Delay representation of a series of 10 000 points generated from the linearly routed logistic equation (see text) in two (upper panel) and three (lower panel) dimensions.

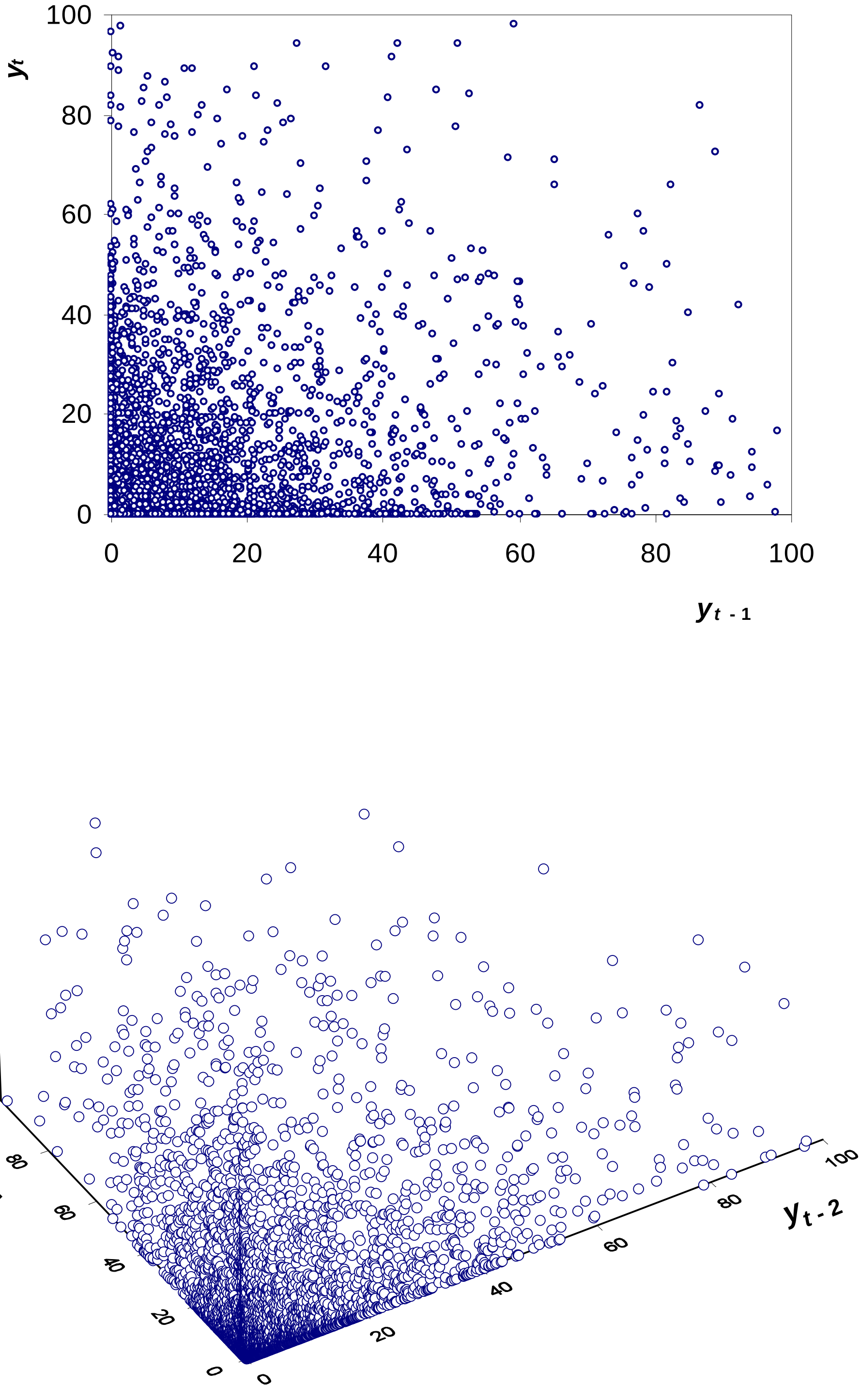

Figure 2 Delay representation of a series of 10 000 daily rainfall depths in two (upper panel) and three (lower panel) dimensions.

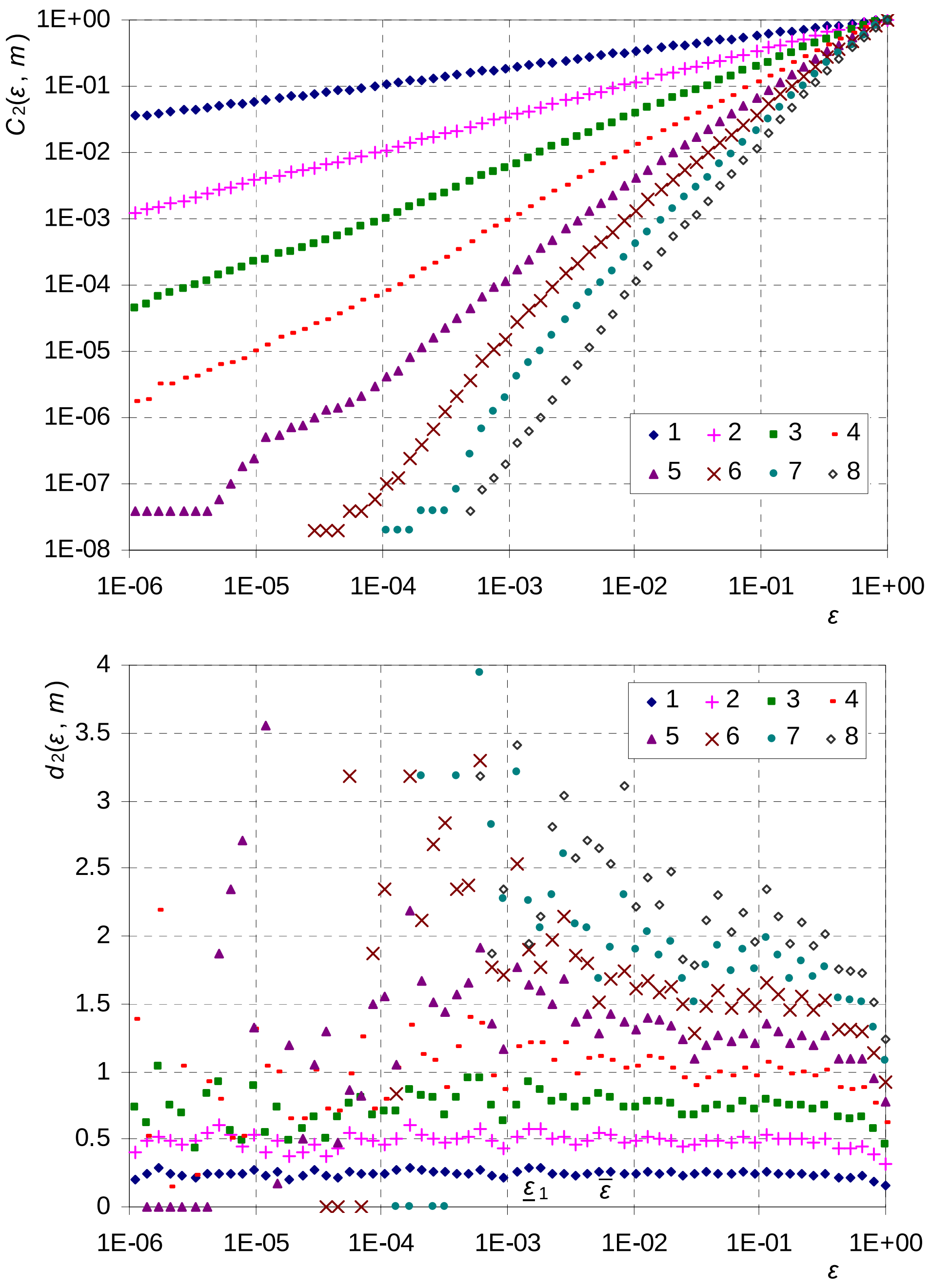

Figure 3 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from a series of 10 000 independent random values with Pareto distribution with exponent 1/8.

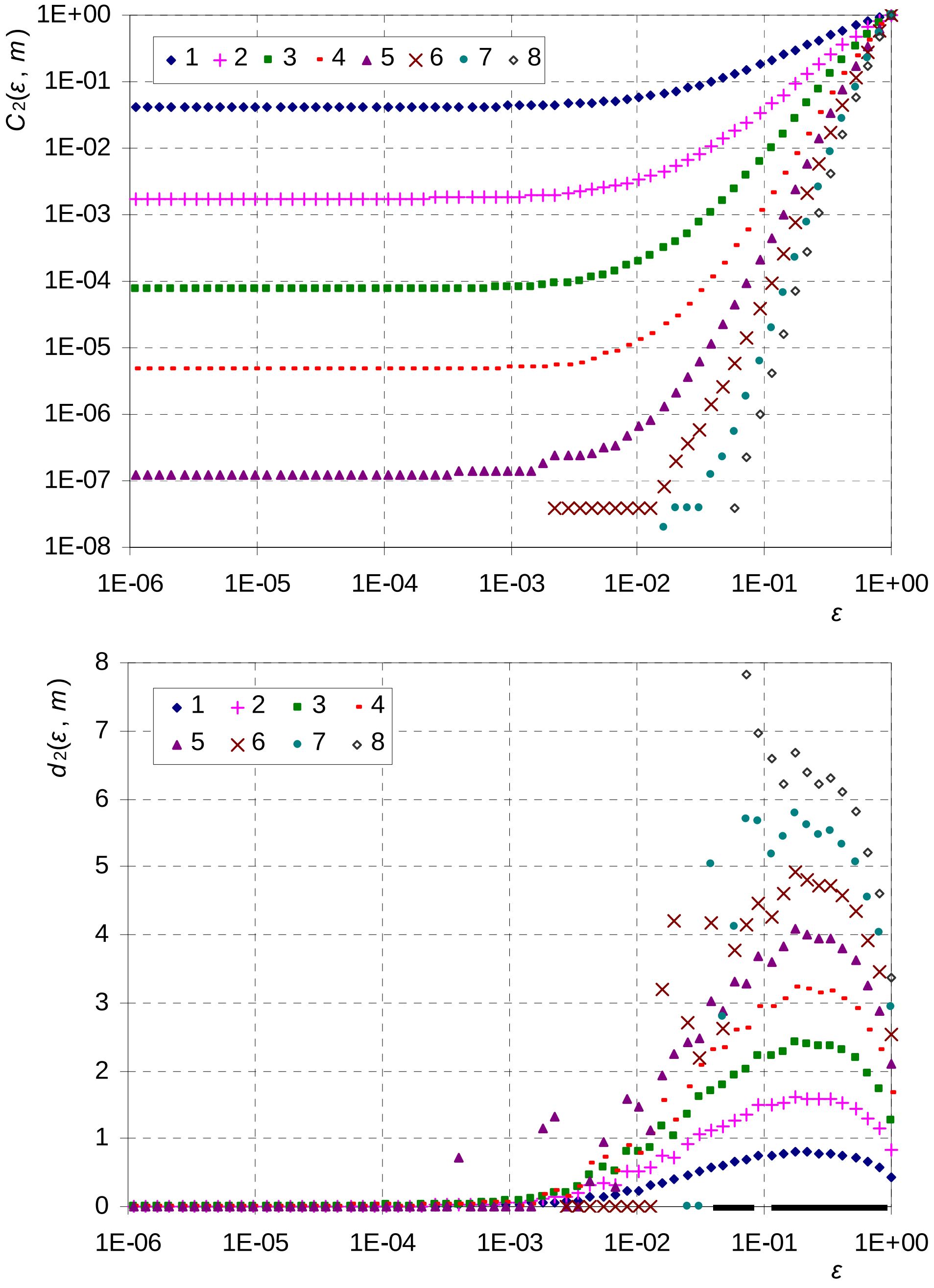

Figure 4 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from a series of 10 000 independent random values, 80% of which are generated from the uniform distribution and the remaining are zeros (located at random).

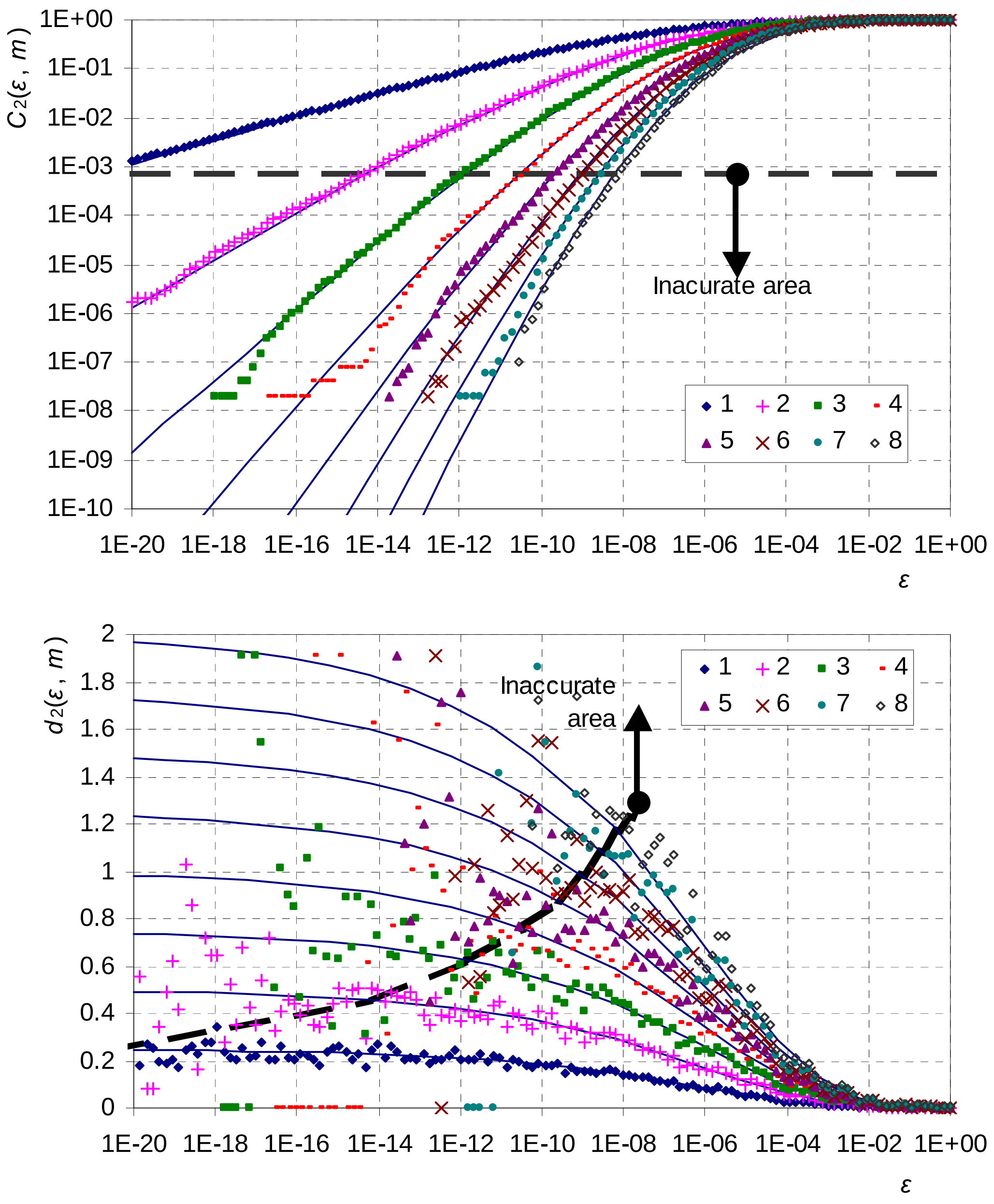

Figure 5 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from a series of 10 000 independent random points from the Weibull distribution with shape parameter 1/8.

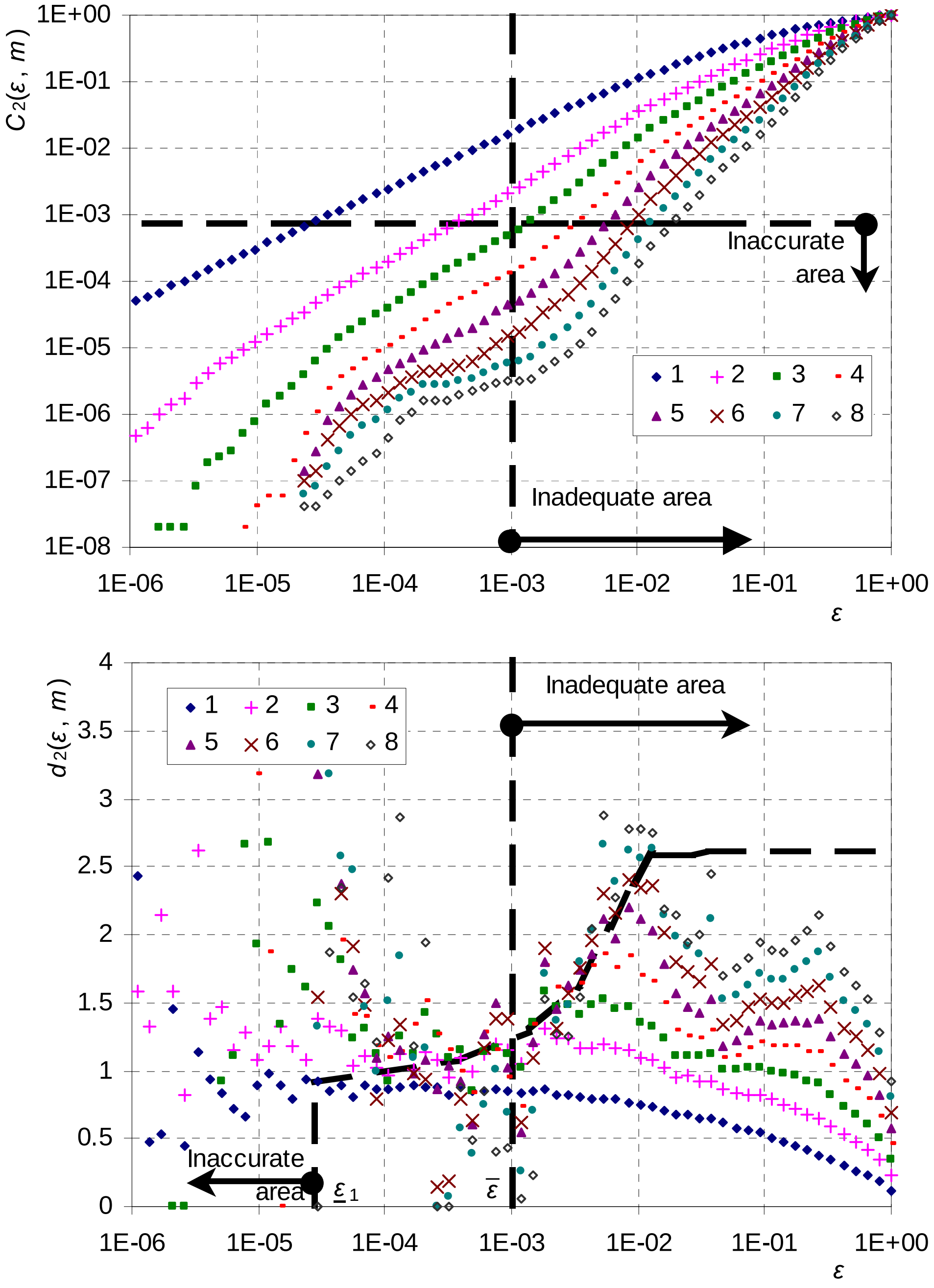

Figure 6 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from a series of 10 000 autocorrelated random values having approximately Pareto distribution with shape parameter 0.44.

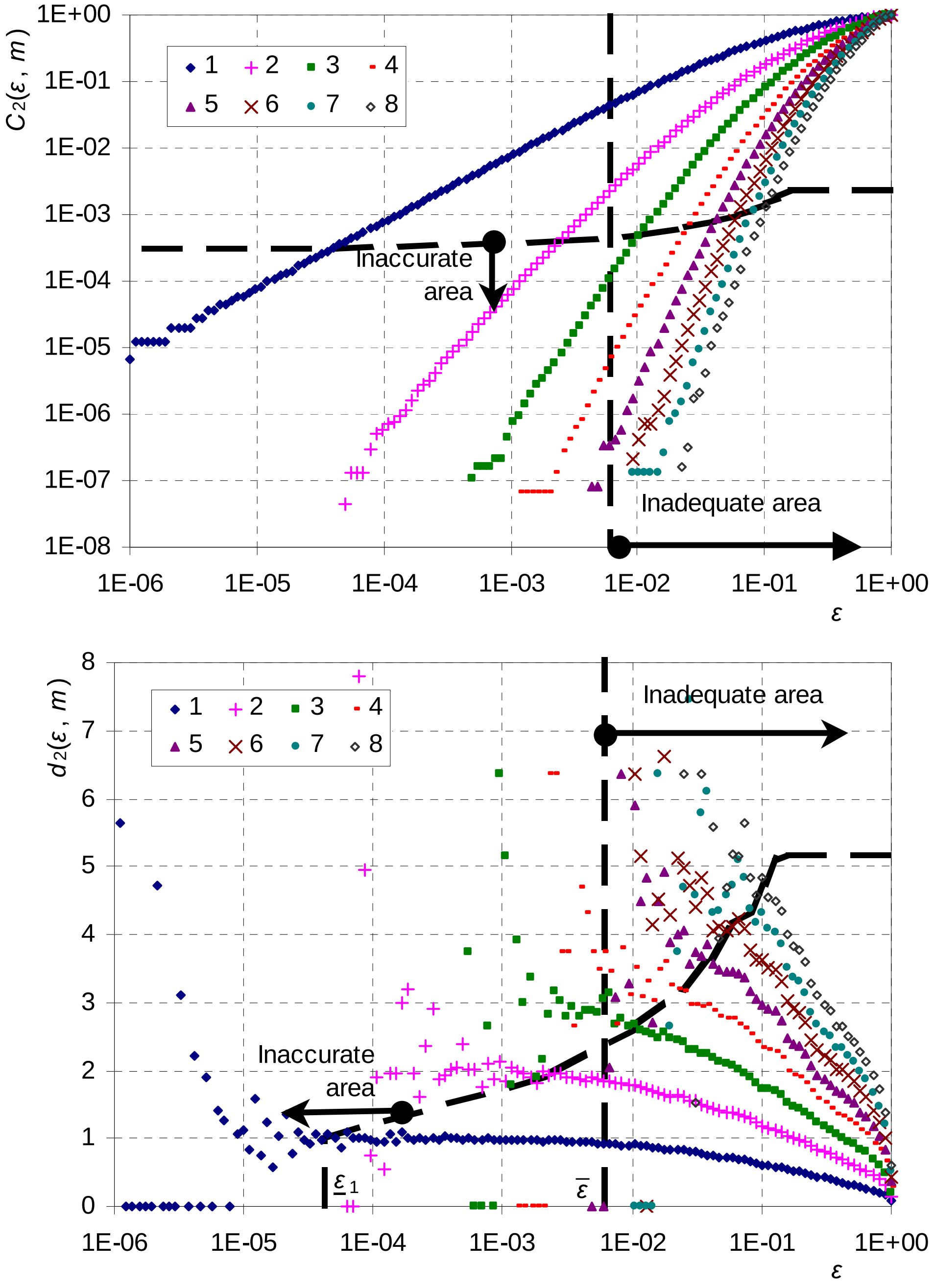

Figure 7 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from the same series as in Figure 6 but excluding points having at least one coordinate smaller than 0.01.

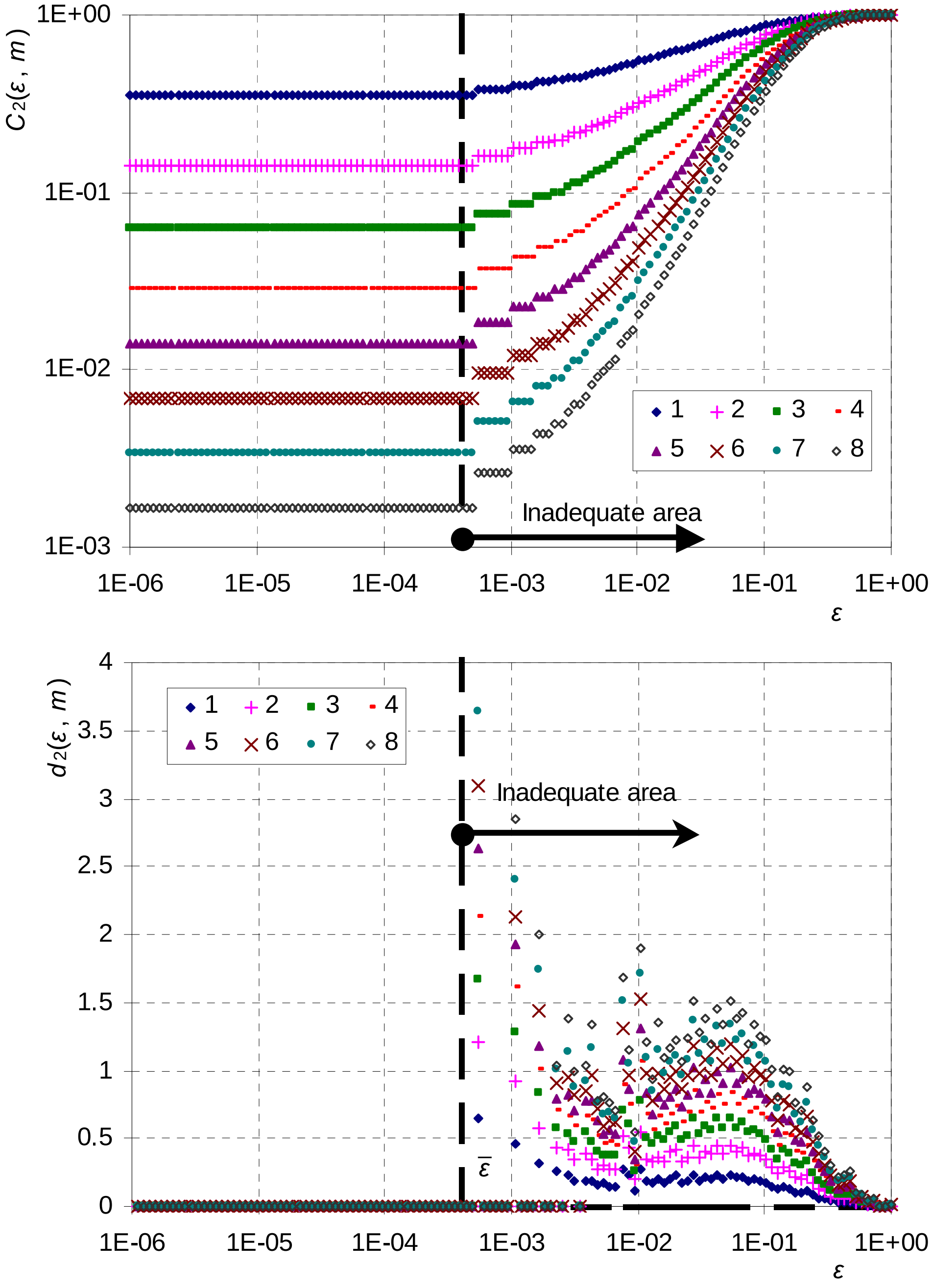

Figure 8 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the daily rainfall series at the Vakari raingage.

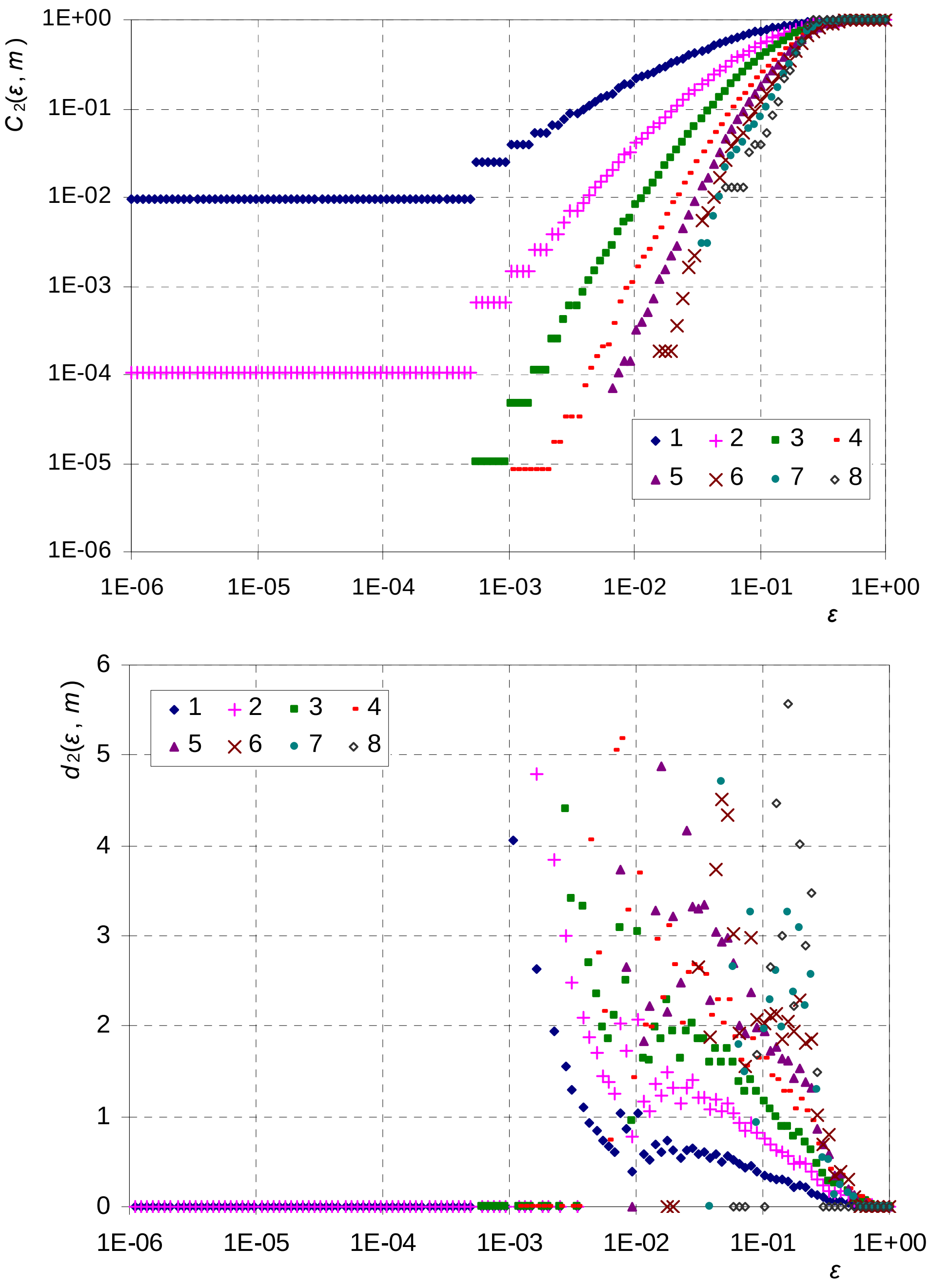

Figure 9 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from the same daily rainfall series as in Figure 8 but excluding points with zero values.

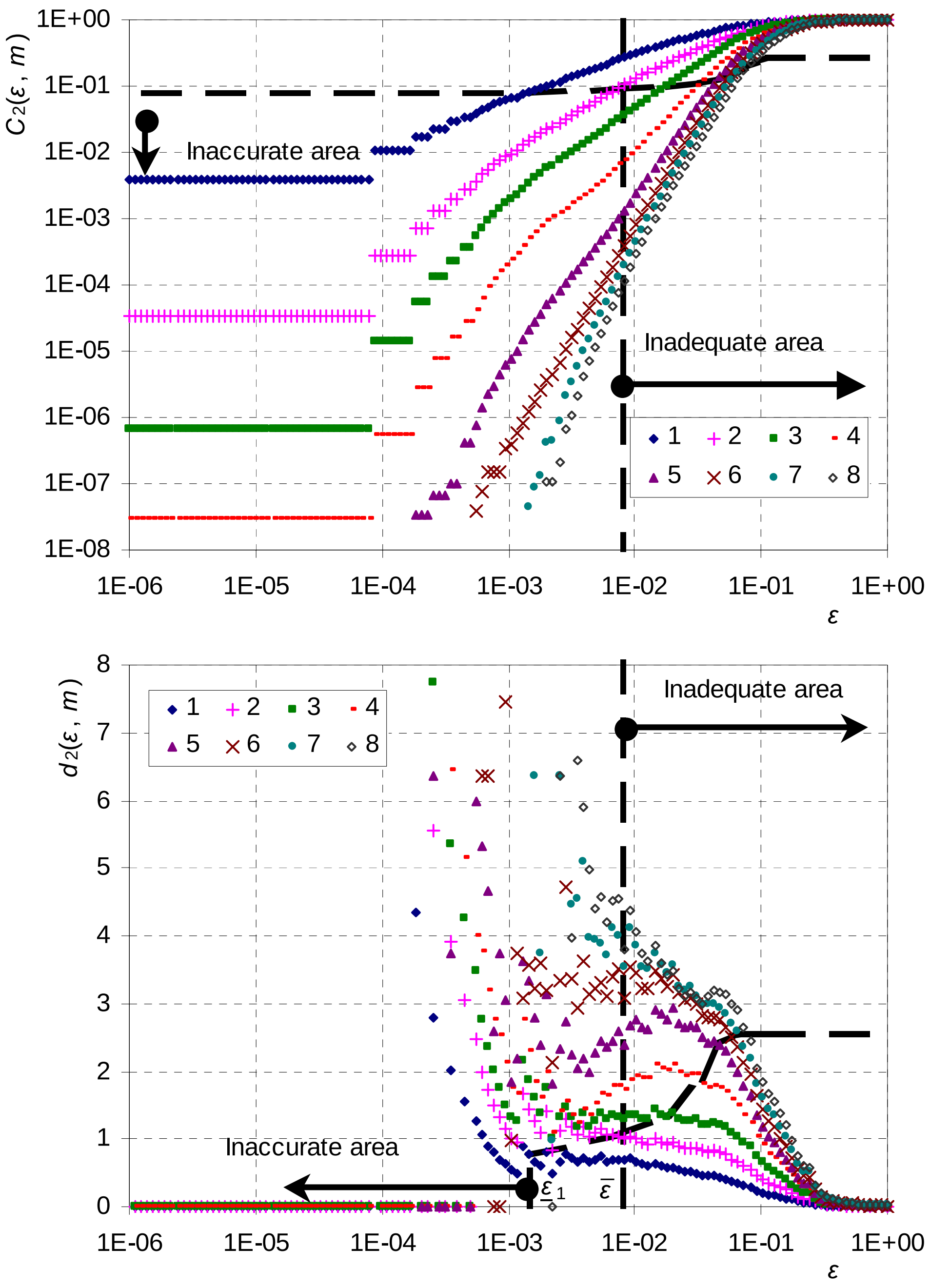

Figure 10 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the fine scale rainfall series at Iowa.

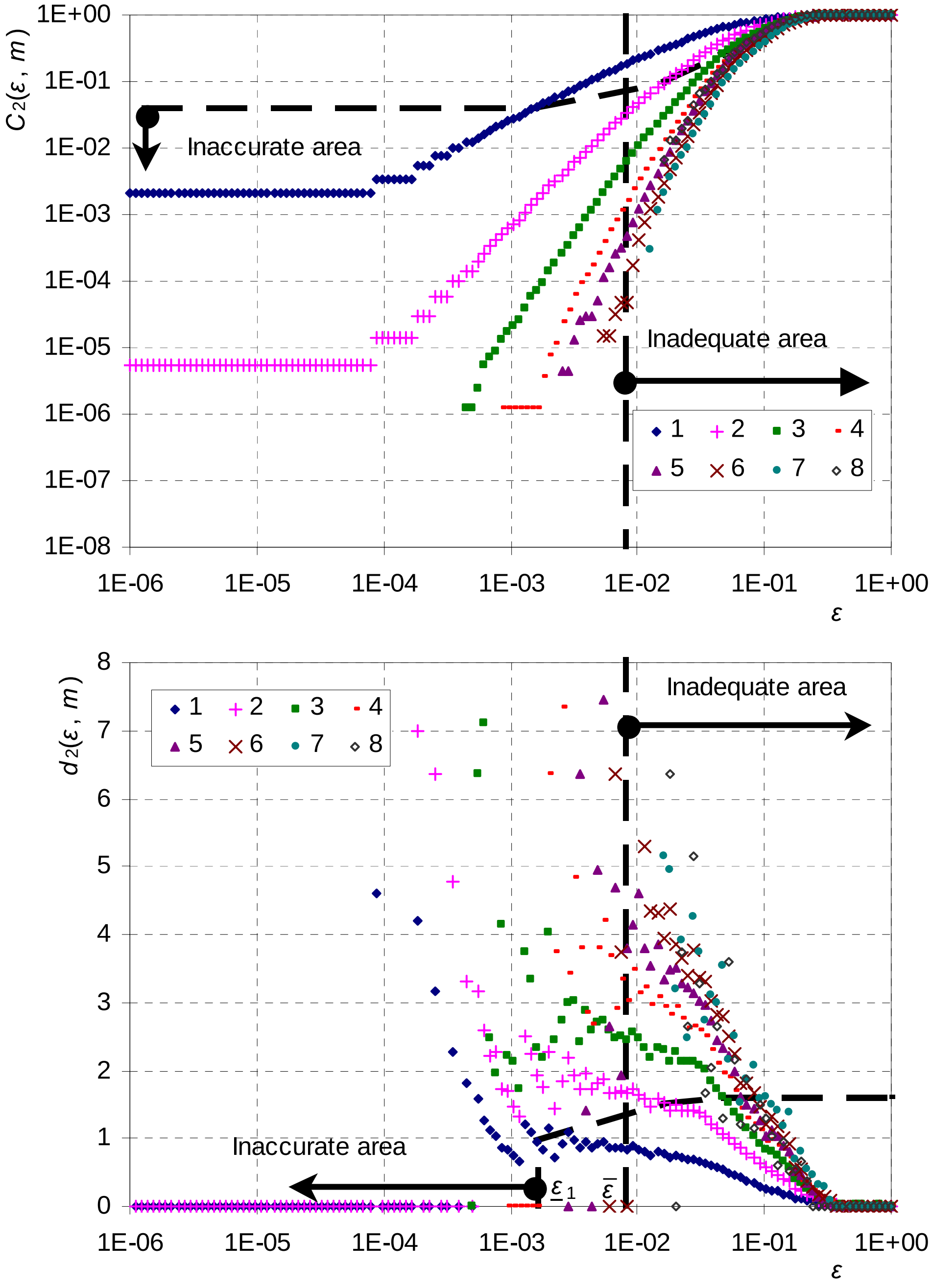

Figure 11 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the fine scale rainfall same series as in Figure 10 but excluding points having at least one coordinate smaller than 0.01.

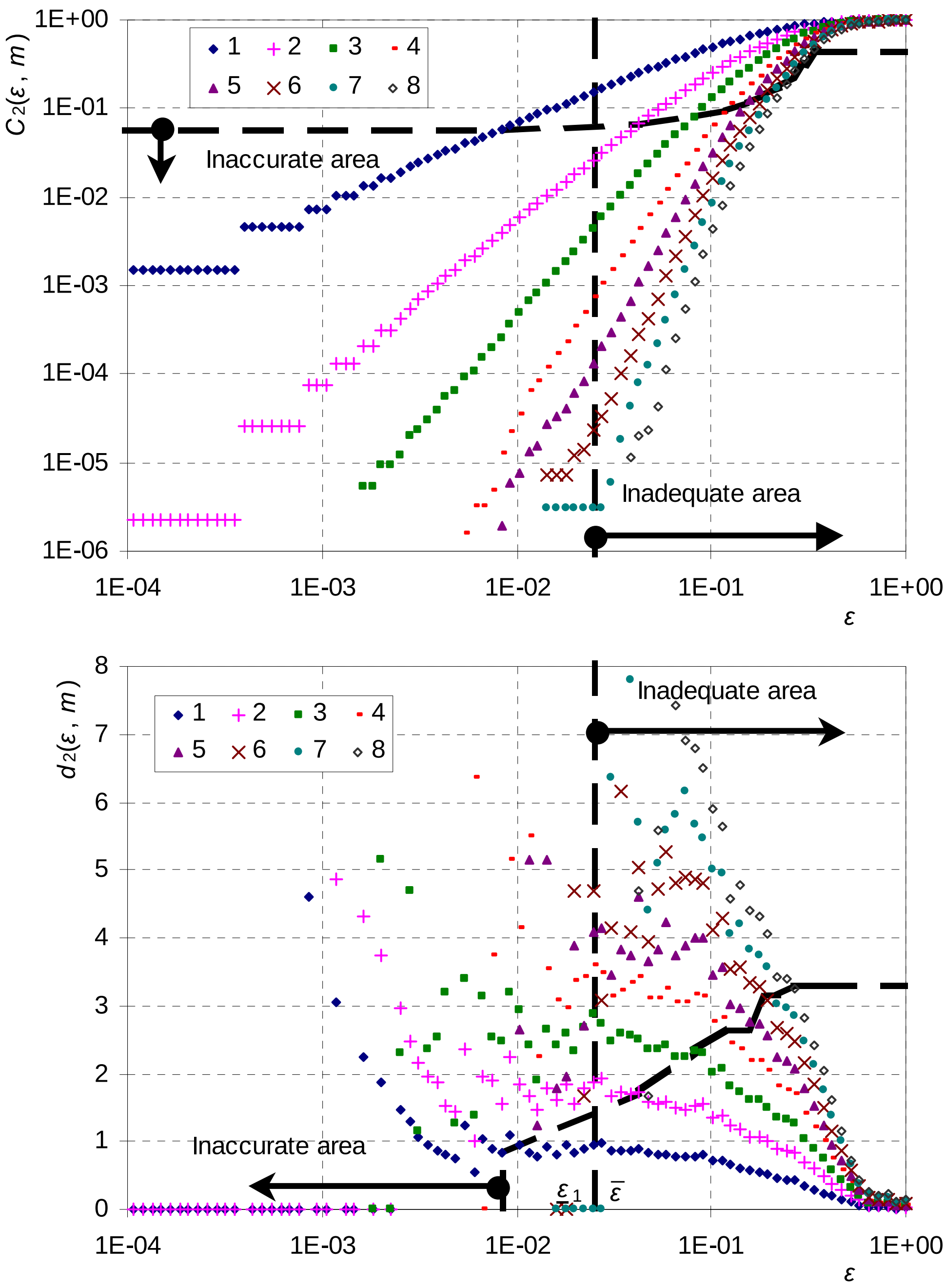

Figure 12 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from the monthly rainfall series at Athens excluding zero points.

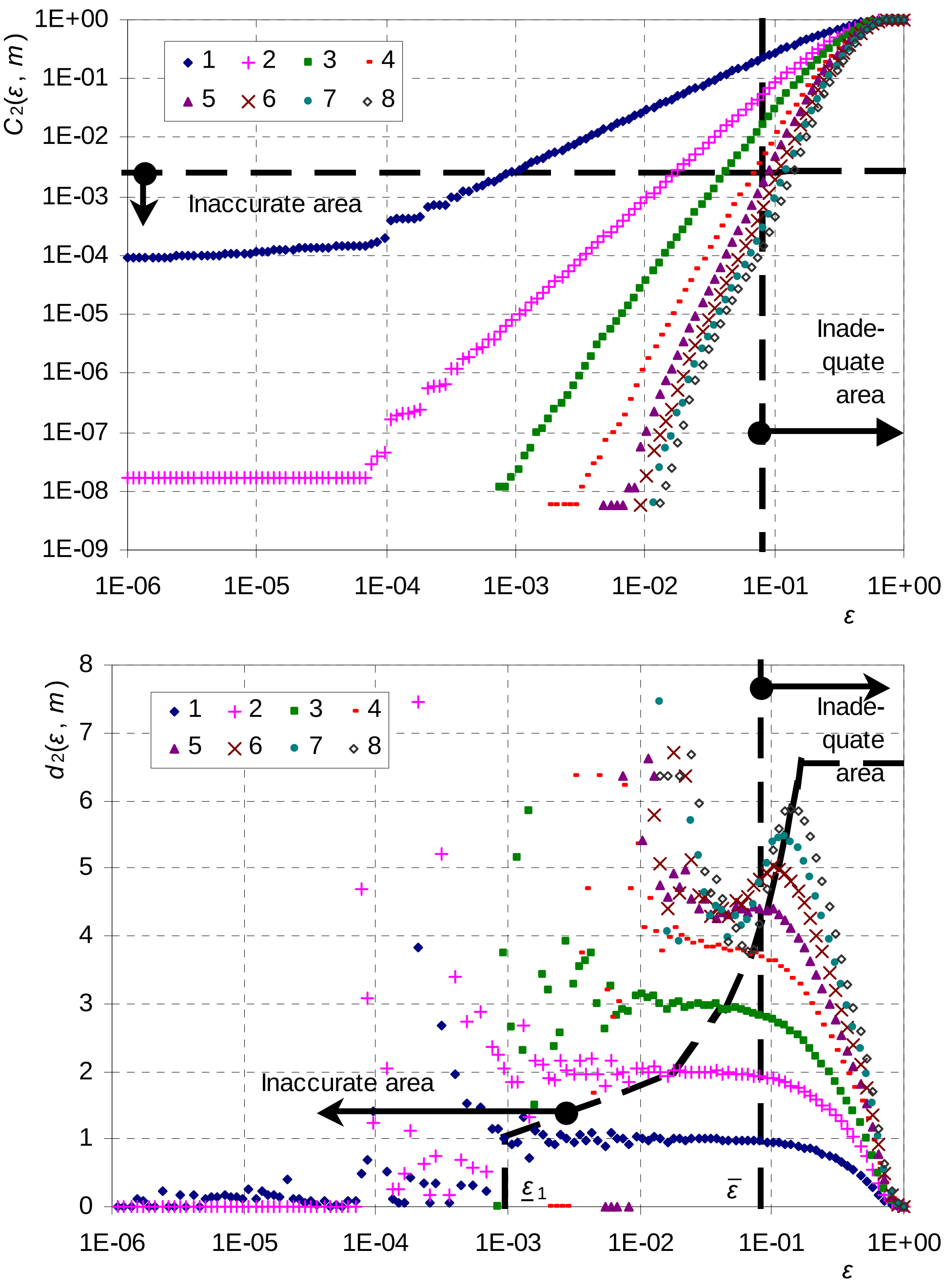

Figure 13 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions $m = 1$ to 8 calculated from the relative humidity series at Athens.

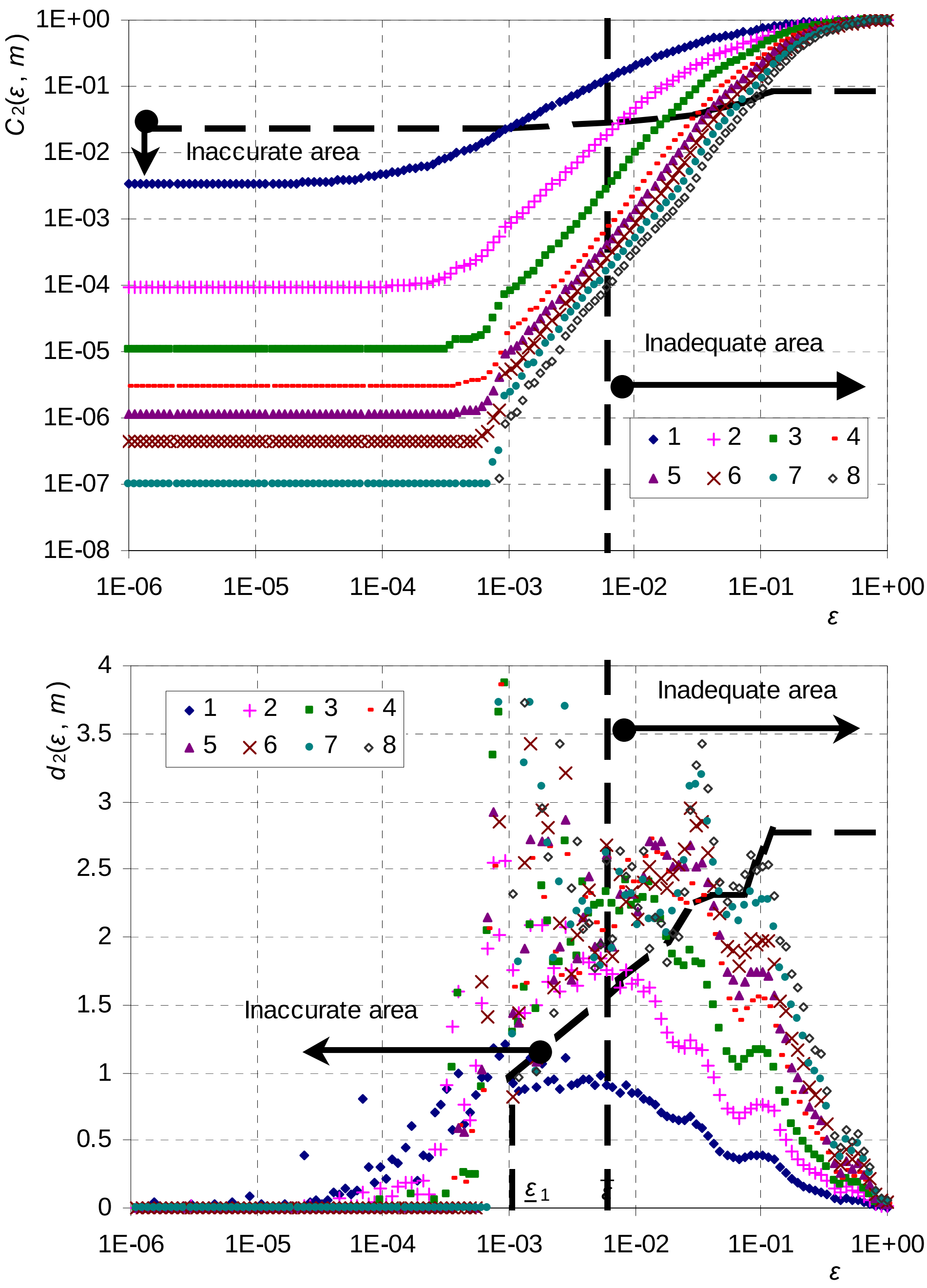

Figure 14 Correlation sums $C_2(\varepsilon, m)$ (upper panel) and their local slopes $d_2(\varepsilon, m)$ (lower panel) versus length scale ε for embedding dimensions m = 1 to 8 calculated from the discharge series at Ali Efenti gage at Pinios River.