DISCUSSION of "Evidence of chaos in the rainfall–runoff process"*

Which chaos in the rainfall–runoff process?

DANIEL SCHERTZER
LMM, Université Pierre et Marie Curie, 4 Place Jussieu, F-75252 Paris Cedex 05, France
and
Météo-France, 1 Quai Branly, F-75007, Paris, France
schertze@ccr.jussieu.fr

IOULIA TCHIGUIRINSKAIA
UMR Sisyphe, LGA, Université Pierre et Marie Curie, 4 Place Jussieu, F-75252 Paris Cedex 05, France
tchigin@biogeodis.jussieu.fr

SHAUN LOVEJOY
Physics Department, McGill University, 3600 University Street, Montreal, Quebec H3A 2T8, Canada
lovejoy@physics.mcgill.ca

PIERRE HUBERT
UMR Sisyphe, Ecole des Mines de Paris, 35 Rue St Honoré, F-77305 Fontainebleau, France
hubert@cig.ensmp.fr

HOCINE BENDJOUDI
UMR Sisyphe, LGA, Université Pierre et Marie Curie, 4 Place Jussieu, F-75252 Paris Cedex 05, France
bdj@biogeodis.jussieu.fr

MICHELE LARCHEVÊQUE
LMM, Université Pierre et Marie Curie, 4 Place Jussieu, F-75252 Paris Cedex 05, France
larchevq@lmm.jussieu.fr

Abstract In the 1980s, there were numerous claims, based on estimates of the correlation dimension, that the variability of various geophysical processes, in particular rainfall, is generated by a low-dimensional deterministic chaos. Due to a recent attempt (Sivakumar *et al.*, 2001) to revive the same approach and with claims of an analogous result for the rainfall–runoff process, we think it is necessary to clarify why this approach can be easily misleading. At the same time, we ask which chaos is involved in the rainfall–runoff process and what are the prospects for its modelling?

Key words rainfall–runoff models; (multi-) fractals; chaotic dynamics; stochastic processes; nonlinear analysis; time series analysis; correlation dimension

* by Bellie Sivakumar, Ronny Berndtsson, Jonas Olsson & Kenji Jinno, *Hydrological Sciences Journal* **46**(1), 131–146.

Open for discussion until 1 August 2002

Quel chaos dans le processus pluie–débit?

Résumé Au cours des années 1980, il a été souvent annoncé, à partir d'estimations de la dimension de corrélation, que la variabilité de différents processus géophysiques, en particulier la pluie, était générée par un chaos déterministe de faible dimension. Du fait d'une récente tentative (Sivakumar *et al.*, 2001) de ressusciter cette approche et l'annonce d'un résultat similaire sur le processus pluie–débit, nous pensons qu'il est nécessaire de clarifier pourquoi cette approche peut être facilement trompeuse. En même temps, nous indiquons quel chaos est en jeu dans le processus pluie–débit et quelles sont les perspectives pour sa modélisation.

Mots clefs modèles pluie–débit; (multi-) fractals; dynamique chaotique; processus stochastiques; analyse nonlinéaire; analyse de séries temporelles; dimension de corrélation

INTRODUCTION

Without doubt the understanding of the dynamics of the rainfall–runoff process constitutes one of the most important and challenging problems in hydrology. Its importance has been emphasized in the prospective of the IAHS decade of "Prediction of Ungauged Bassins".

Unfortunately, in our opinion, the paper by Sivakumar *et al.* (2001) does not seem to contribute to a clarification of these issues and may lead to erroneous conclusions. This paper, as well as several similar ones written or co-authored by the same lead author (Sivakumar *et al.*, 1999a,b; Sivakumar, 2000), tries to show with the help of the correlation dimension that hydrological processes have a low-dimensionality (more precisely of the order 6–7), and therefore should fall into the category of (low-dimensional) "deterministic chaos". One can recall that it has been known for the last 15 years that such an approach may be misleading.

Since Sivakumar *et al.* (2001) repudiated this knowledge by calling it "belief", we will emphasize the well-known fact that an empirical low value of dimension can easily be spurious. It can be an artefact of the finite size of the data set, rather than a reliable estimate of the dimensionality of the process, or it can result from the stochastic nature of the process.

With the help of a synthetic series generated by a stochastic process related to the rainfall process, we give a concrete example. We conclude with a discussion on the type of chaos that is involved in the rainfall–runoff process.

WHAT IS CHAOS?

Origin of the chaos concept

Chaos is one of the oldest concepts and paradigms. Its origin could be traced back at least to Greek mythology. In this framework, Chaos was the disordered world that preceded Cosmos that is our present ordered world. Therefore, in its wide sense, chaos merely refers to some kind of disorder. And indeed, it has been used for rather distinct types of disorder. For instance, Wiener (1938) called "pure chaos" the Brownian motion, which he mathematically formalized. More recently, Kahane (1995) referred to "Lévy chaos" as the extension of this pure chaos to motion, defined with the help of the Lévy variables, which broadly generalize the Gaussian variables.

Deterministic chaos

During the last 30 years, chaos took a much more a restrictive meaning, since it became understood as a shorthand for "deterministic chaos". The latter denotes the disorder generated by deterministic dynamical systems. Here the adjective "deterministic" means that the equations do not contain any noise source, and that, in general (but not always), the existence and uniqueness of solutions are mathematically assured. The main defining feature of such a chaotic system is the sensitive dependence of solutions on the initial conditions (and/or boundary conditions for partial differential systems). The prototype example is the celebrated Lorenz model (Lorenz, 1963), which was introduced as a mathematical caricature of atmospheric convection and has the lowest possible dimensionality, i.e. three, for chaotic differential systems.

This chaos was initially viewed as a mathematical curiosity, e.g. the Lorenz model attracted little attention until the development of *nonlinear time series analysis*, about 20 years ago, with the pioneering work by Packard *et al.* (1980) and the seminal paper by Takens (1980). The core of this approach corresponds to the possible "reconstruction" of the phase space from (discrete) time series $X_n = h[x(n\Delta t)]$ for a given scalar observable h of a (vector-valued) trajectory $\boldsymbol{x}(t)$.

This can be achieved with the help of the m-dimensional "embedding space", E_m which is spanned by delay vectors $\boldsymbol{Y}_n = (X_n, X_{n-\tau}, X_{n-2\tau}, ..., X_{n-(m-1)\tau})$, for any given (integer) time delay τ. More precisely, a very interesting extension of the classical Witney embedding theorem (for integer dimensional manifolds) was obtained. Indeed, Takens (1980) and Sauer *et al.* (1991) demonstrated that if the trajectory $\boldsymbol{x}(t)$ is confined on an invariant set A, with box dimension D, then there is a unique and invertible smooth map from A to the delay embedding space E_m, for $m > 2D$. As a consequence, the dynamics can be represented in the m-dimensional delay embedding space E_m, i.e. with m independent variables.

On the theoretical level, this theorem is extremely appealing, since one needs only a given scalar observable h at discrete times ($n\Delta t$) in order to get the dynamics of a vector-valued process $\boldsymbol{x}(t)$ for continuous time (t). However, the underlying hypothesis is rather demanding (see comment below). Furthermore, for practical reasons (discussed below), its applications were primarily restricted to very low-dimensional systems, say a dimension not much higher than 5, whereas the mathematical theory does not face such a limitation. Indeed, the main mathematical difficulties are related to the transition from finite dimensional systems to infinite systems, e.g. partial differential systems, and the corresponding disorder is often called "spatially extended chaos", which is a rather new field.

This practical restriction to a very narrow range of (low) dimensions had the consequence that many practitioners believed not only that low dimensionality is a requisite of deterministic chaos, but also that its empirical evidence is "an indication of the possible existence of chaos". This corresponds to a misinterpretation of the embedding theorem that is discussed below in more detail. Furthermore, it turns out that nonlinear time series analysis techniques are inherently incapable of distinguishing between low-dimensional deterministic systems and high-dimensional stochastic systems (see "stochastic chaos" below).

Stochastic chaos

Fortunately, in spite of the big impetus of the deterministic chaos theories, there have also been important developments in stochastic modelling. Some of the reasons need to be reviewed, since Sivakumar *et al.* (2001) seem to rule out a possible relevance of stochastic modelling to hydrological processes.

Firstly, contrary to deterministic chaos, stochastic models can easily simulate spatially extended systems. In particular, this is notable for turbulent fields. Indeed, it has become widely recognized that models with a small number of degrees of freedom are inadequate to model turbulence, except in the low Reynolds number regime below the transition to turbulence. For instance, Ruelle (1989) pointed out "the gap dividing simple chaotic systems and fully developed turbulence". Indeed, a qualitative new understanding of the fundamental problem of the intermittency of turbulence, i.e. the fact that only a relatively small fraction of the enormous number of degrees of freedom are effectively active, was obtained with the help of a rather new type of stochastic process. Inspired by the cascade paradigm, the latter are multiplicative rather than additive, as for the classical random walks (and associated diffusive processes). The term "multiplicative chaos" was indeed proposed by Kahane (1995), although the name "multifractal" became much more widely used. In any event, stochastic multifractal processes have been increasingly considered in geophysics (Schertzer & Lovejoy, 1991; Schertzer *et al.*, 1997), in particular in hydrology (Lovejoy & Schertzer, 1995).

Secondly, the deterministic features of a low-dimensional process can be easily concealed by contamination by a weak noise. The origin of this contamination is not only measurement noise (which could be handled in most cases by specific noise reduction procedures), but also interactions with other scales of the same process or with other processes. In other words, deterministic chaos has the very serious drawback of not being "robust"!

There has therefore been a renewed interest in stochastic differential equations, i.e. dynamic equations that include a noise source term. For instance, it has been known for decades that a local Gaussian perturbation leads to the classical Fokker-Planck equation (e.g. Van Kampen, 1981). But it was only recently demonstrated that strongly non-Gaussian perturbations lead to a broad "fractional" generalization of the Fokker-Planck equation (e.g. Schertzer *et al.*, 2001), i.e. a kinematic equation for the probability, which involves fractional derivatives. Furthermore, there are expectations that this type of equation can generate multifractal fields.

Furthermore, multifractal processes help to combine stochastics with dynamics, since they are based on both physics and statistics. This is particularly the case for hydrology, since, as emphasized by Hubert *et al.* (1993), it reconciles two opposing views on extreme precipitation, the "extreme maximum precipitation" (PMP) and probability approaches (based on frequency analyses). Indeed, multiplicative cascades account for turbulent processes resulting from nonlinear interactions between different scales and fields and respect a basic symmetry of the nonlinear generating equations, i.e. scale invariance. Furthermore, the details of a multifractal process are theoretically defined by the statistical distribution of the singularities of the generating equations. For applications, there are various reliable techniques to extract this distribution from empirical time and/or space series.

Finally, there is no obvious reason that processes should be run by deterministic equations rather than by stochastic equations, since the former are merely particular cases of the latter. Therefore, one can question the deterministic reductionism that is rather ubiquitous in the natural sciences, i.e. a tendency to look at deterministic systems as if they were the only ones providing causality. It is rather important to appreciate that this tendency corresponds to philosophical bias, rather than to an objective rationality and that, behind the notion of deterministic chaos, there could be a possible resurgence of some former restrictive notions of determinism (see Lovejoy & Schertzer, 1998).

THE INTEREST, LIMITATIONS AND PITFALLS OF THE CORRELATION DIMENSION

A straightforward method

Grassberger & Procaccia (1983) introduced a rather efficient algorithm for dimension estimation, which became extremely popular and therefore available in several numerical packages (e.g. Press *et al.*, 1986). One fundamental reason of its popularity is that, in contrast to a box counting algorithm, the embedding delay space is only implicit rather than explicit, since only the distance between pairs of delay vectors is required. As a consequence, much larger embedding dimensions m can be numerically explored than for a box counting dimension algorithm. Indeed, the correlation dimension D_2 is defined as the scaling exponent of the average number of delay vectors in a sphere of radius r centred on one of them, i.e.:

$$< N(m,r) > \;\propto r^{D_2(m)} \tag{1}$$

and this average number is precisely defined with the help of the correlation sum:

$$< N(m,r) > \equiv C(m,r) = \frac{2}{(N-T)(N-T-1)} \sum_{i<j-T} H(l - |s_i - s_j|) \tag{2}$$

where N is the number of data points, H is the Heaviside function ($x > 0 : H(x) = 1; x \le 0 : H(x) = 0$) and $T = 0$ is the Theiler window parameter (Theiler, 1986), which is optionally introduced to suppress trivial pairs having too close time indices.

For large embedding dimension m, the estimates $D_2(m)$ should converge towards the theoretical value D_2. In other words, the curves $C(m,r)$ *vs* r in a log-log plot should lie on the same straight line (having a slope D_2), at least over a given range of r. This is fairly straightforward and it has been extensively used for many different data sets in physics and geophysics. In this respect, Fig. 3(a) and (b) of the paper by Sivakumar *et al.* (2001) is very suggestive. This is much the same for our Fig. 1, which is discussed below.

Theoretical limitations

In fact the correlation dimension D_2 and the box dimension D_0 are two special cases of the infinite hierarchy of so-called Renyi dimensions (Grassberger, 1983) that characterize the multifractal behaviour of a strange attractor. They are in general distinct and only related by the following inequality:

$$D_0 \geq D_2 \tag{3}$$

Therefore, D_2 yields only a lower bound of the box dimension, whereas the latter is the dimension required for the embedding theorem. Therefore, even for deterministic systems, D_2 may largely underestimate the dimensionality of the dynamics.

Contrary to a frequent misinterpretation of this theorem, it is important to note that the theorem *hypothesizes* that the dynamics are deterministic. Therefore it does not draw any conclusion from the mere determination of a low dimensionality. There are obvious reasons for this. Indeed, there are well known stochastic processes having a low dimensionality. The most celebrated one is Brownian motion (Osborne & Provenzale, 1989; Theiler, 1991), since this additive process is known to have a box-counting dimension $D_2(m) = 2$ for any m!

Figure 1 corresponds to a correlation-dimension analysis of a synthetic series (displayed in Fig. 2) simulated with the help of a stochastic cascade process. This *nonlinear* type of process was chosen, because it has been often invoked for rainfall data analysis and simulation, as discussed above. However, to keep this process as simple as possible, so that any curious reader could reproduce it, a discrete (in scale) "lognormal" cascade was chosen, whereas numerous studies show that the rainfall cascade is continuous in scale and rather "log-Lévy". Nevertheless, independently of the details of such a cascade process, it has a very large dimensional phase space (infinite dimensional if the cascade process proceeds down to an infinitesimal "inner" scale) and an infinite dimensional probability space. In spite of the large dimension of the space, the correlation dimension yields a low finite estimate D_2. In our precise example, there are 12 cascade steps, the number of data points is $N = 4096$ and the mean fractality of the cascade is $C_1 = 0.03$, and $D_2 \approx 2.7$ was obtained numerically.

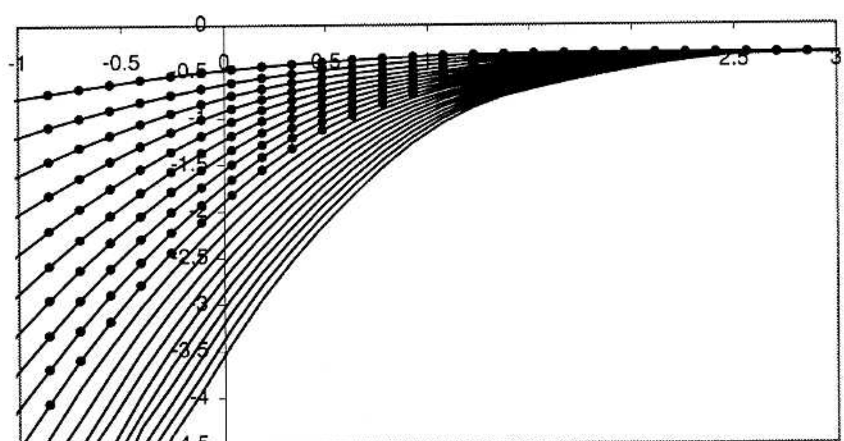

Fig. 1 Log $C(r,m)$ *vs* log$r$ for the synthetic time series displayed in Fig. 2, for embedding dimensions $m = 1, 2, ..., 20$ (top to bottom). The corresponding exponent $D_2(m)$ (estimated in the scaling range) converges towards a low dimension $D_2 \approx 2.7$, whereas the stochastic process has an infinite dimensional probability space and a very large dimensional phase space.

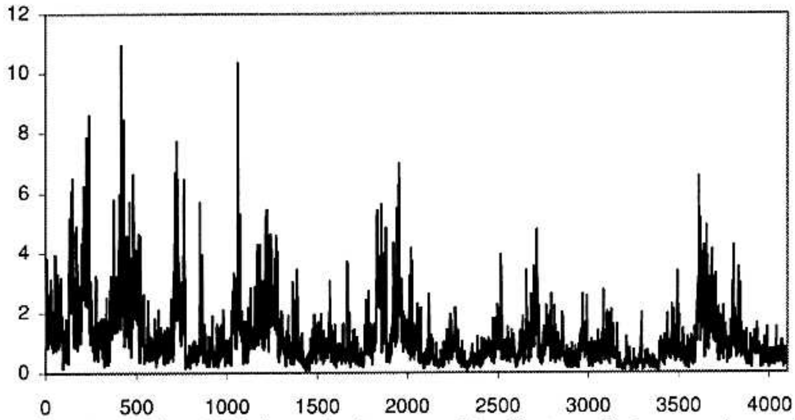

Fig. 2 Time series (4096 data points) generated by 12 steps of a lognormal cascade, which is defined by its mean fractality $C_1 = 0.03$.

Pitfalls

The ease of the correlation-dimension method is also a source of pitfalls. Indeed, the fact that the embedding space is only implicit may lead one to forget an apparently straightforward, but essential, question: What is really being estimated when the embedding dimension becomes larger and larger?

The absence of problems with numerics should not hide the fact that there could be another obvious problem. Indeed, any empirical analysis is performed on a finitely sized sample, and the confinement of empirical points measured by $D_2(m)$ to a small fraction of the embedding space could be due to the limited number of points rather than the dynamics! In other words, instead of measuring an effect of the dynamics, one is merely evaluating an artefact of the finite sample size! One simple way (e.g. Grassberger, 1986) of estimating the latter is to come back to the box counting dimension. Indeed, according to this notion, the number of points spread homogeneously over a fractal set of dimension D and on a scale ratio λ scales like:

$$N(\lambda) \propto \lambda^{D} \tag{4}$$

Assuming that a decade in scales is a minimum to demonstrate a scaling behaviour, one obtains the celebrated rule of thumb that the minimal number of points to estimate a dimension D is:

$$N_D \approx 10^{D} \tag{5}$$

This rule explained many of the unusual results obtained with the help of the correlation dimension. In their pioneering and influential paper, Nicolis & Nicolis (1984) used 500 values from the isotope record of a deep-sea core to conclude that $D_2 \approx 3.1$. According to this estimate, one might be able to create a model predicting climatic changes of the last million years with only 7–8 independent variables. Grassberger (1986) discussed these results. In fact, the 500 values were obtained by interpolation of only 184 actual measurements. Applying the rule of thumb (equation (5)), one obtains that the maximal dimension, which could be safely estimated, is of the order $D_2 \approx 2.3$.

Grassberger (1986) concluded that it is difficult to distinguish these data from those of a random signal series.

Nerenberg & Essex (1990) and Essex (1991) refined somewhat the rule of thumb and obtained a slightly more optimistic estimate of the number points required to obtain a reliable dimension:

$$N_D \approx 10^{2+\ 0.4D} \qquad (6)$$

which corresponds to the introduction of an explicit prefactor in equation (5), and to considering a smaller range of scale ($\lambda \approx 2.5$).

Let us note that in the case of the (stochastic) multiplicative cascade displayed in Fig. 2, the numerical estimate of the corresponding correlation dimension (Fig. 1) $D_2 \approx 2.7$ is rather reliable, since $10^{2.7} \approx 500$, whereas we have $4096 \approx 10^{3.6}$ points.

Chaos and the rainfall–runoff process

To come back to the question of chaos and of the rainfall–runoff process: no-one will question the erratic nature of this process, and therefore its chaotic nature in the wide sense. However, many objections will be raised if chaos should be understood in the narrow sense of deterministic chaos, furthermore with a low dimensionality, say 5–6. Indeed, it flies in the face of common sense that rainfall could be predicted with twelve independent variables. At the very least, one would rather think that some set of partial differential equations is required!

The rule of thumb (equation (5)) yields an upper bound for a maximal reliable estimate of the order $D \approx 3.2$, whereas Sivakumar *et al.* (2001) claim to have reliable estimates of $D \approx 5$–6. In other words, they have only 1% of the necessary data set. Nevertheless, the authors are aware of the "optimistic" evaluation (equation (6)) by Nerenberg & Essex (1990) of the necessary number of data points in order to obtain their estimates in a reliable manner. Corresponding to this estimate, their data contain only about 10% of what would be necessary. They merely repudiate any evaluation of this type, by calling it a "belief"! They claim that "for a particular size, the number of reconstructed vectors may not differ much whether an embedding dimension of, for example, 4 or 10 is used, and, therefore the dimension estimate may not be influenced much". This statement reflects a misunderstanding of the arguments discussed in the previous section. Indeed, the estimate of the necessary number of points (equation (4)) is related to the expected dimension of the attractor as well as to the scale ratio of the scaling range, and *not* directly to the embedding dimension! Furthermore, they claim that the only important question is to obtain "a large scaling region", which indeed corresponds to the second factor, but not the only one, of these arguments. Unfortunately, they forget to evaluate how narrow is their scaling region. For instance, their Fig. 3(a) displays a scaling range that seems to be of the order of $\lambda \approx 3$. Let us add that we do not understand why they drop a factor of 2 in the estimate of the necessary number of independent variables, i.e. considering it as D instead of $2D$. In any case, Sivakumar *et al.* (2001) were unable to substantiate their claim of "an indication of low-dimensional chaotic behaviour" in the rainfall–runoff process.

CONCLUSIONS

Some time ago, low-dimensional deterministic chaos had been very helpful in order to better understand the limitations of classical methods in analysing and modelling complex systems, in particular in hydrology. This was achieved with the help of an apparently simple caricature of a complex system (e.g. the Lorenz model corresponds to the truncation of the convection of the first three Fourier modes of convection) leading nevertheless to nontrivial behaviours.

However, in the name of a mathematical theorem—in fact a fundamental misinterpretation of this theorem—there had been an awkward tendency to attempt to reduce complex systems to their low-dimensional caricatures. This tendency was reinforced by the apparent success of a fairly straightforward algorithm to estimate rather low dimensionality for various complex systems. However, many of these estimates may easily turn out to be spurious, either because of sample size limitation or of the stochastic nature of the process. The paper by Sivakumar *et al.* (2001) rather corresponds to a late confirmation of this dead end.

In this discussion it has been pointed out that the chaos of spatially extended systems, which include hydrological systems, may require approaches dealing with a very large degree of freedom and that some asymptotic behaviours correspond to infinite numbers. It has also been pointed out that progress in that direction might result from an original blending of stochastics and dynamics.

REFERENCES

Essex, C. (1991) Correlation dimension and data sample size. In: *Non-Linear Variability in Geophysics, Scaling and Fractals* (ed. by D. Schertzer & S. Lovejoy). Kluwer, Dordrecht, The Netherlands.

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

Grassberger, P. (1986) Are there really climate attractors? *Nature* **322**, 609–612.

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

Hubert, P., Tessier, Y., Ladoy, Ph., Lovejoy, S., Schertzer, D., Carbonnel, J. P., Violette, S., Desurosne, I. & Schmitt, F. (1993) Multifractals and extreme rainfall events. *Geophys. Res. Lett.* **20**(10), 931–934.

Kahane, J. P. (1995) Definition of stable laws, infinitely divisible laws, and Lévy processes. In: *Lévy Flights and Related Phenomena in Physics* (ed. by M. Shlesinger, G. Zaslavsky & U. Frisch), 99–109. Springer-Verlag, Berlin, Germany.

Lovejoy, S. & Schertzer, D. (1995) Multifractals and rain. In: *New Uncertainty Concepts in Hydrology and Water Resources* (ed. by Z. W. Kundzewicz), 62–103. Cambridge University Press, Cambridge, UK.

Lovejoy, S. & Schertzer, D. (1998) Stochastic chaos, scale invariance, multifractals and our turbulent atmosphere. In: *ECO-TEC: Architecture of the In-between* (ed. by A. Marras), 80–99. Storefront Book Series, copublished with Princeton Architectural Press, Princeton, USA.

Lorenz, E. N. (1963) Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141.

Nerenberg, M. A. H. & Essex, C. (1990) Correlation dimension and systematic geometric effects. *Phys. Rev. Lett.* **A42**(12), 7065–7074.

Nicolis, C. & Nicolis, G. (1984) Is there a climate attractor? *Nature* **311**, 529–533.

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

Packard, N. H., Crutchfield, J. P., Farmer, J. D. & Shaw, R. S. (1980) Geometry from a time series. *Phys. Rev. Lett.* **45**(9), 712–716.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1986) *Numerical Recipes, The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK.

Ruelle, D. (1989) *Chaotic Evolution and Strange Attractors*. Cambridge University Press, Cambridge, UK.

Sauer, T., Yorke, J. & Casdagli, M. (1991) Embedology. *J. Statist. Phys.* **65**, 579.

Schertzer, D. & Lovejoy, S. (eds) (1991) *Non-Linear Variability in Geophysics, Scaling and Fractals*. Kluwer, Dordrecht, The Netherlands.

Schertzer, D., Lovejoy, S., Schmitt, F., Chigirinskaya, Y. & Marsan, D. (1997) Multifractal cascade dynamics and turbulent intermittency. *Fractals* **5**(3), 427–471.

Schertzer, D., Larchevêque, M., Duan, J., Yanovsky, V. V. & Lovejoy, S. (2001) Fractional Fokker–Planck equation for nonlinear stochastic differential equations driven by non-Gaussian Lévy stable noises. *J. Math. Phys.* **41**, 12.

Sivakumar, B. (2000) Chaos theory in hydrology: important issues and interpretations. *J. Hydrol.* **227**(1,E4), 1–20.

Sivakumar, B., Liong, S. Y., Liaw, C. Y. & Phoon, K. K. (1999a) Singapore rainfall behaviour: chaotic? *J. Hydrol. Engng ASCE* **4**(1), 38–48.

Sivakumar, B., Phoon, K. K., Liong, S. Y. & Liaw, C. Y. (1999b) A systematic approach to noise reduction in chaotic hydrological time series. *J. Hydrol.* **219**(3/4), 103–135.

Sivakumar, B., Berndtsson, R., Olsson, J. & Jinno, K. (2001) Evidence of chaos in the rainfall–runoff process. *Hydrol. Sci. J.* **46**(1), 131–146.

Takens, F. (1980) Detecting strange attractors in turbulence. In: *Dynamical Systems and Turbulence, Lecture Notes in Mathematics 898* (ed. by D. A. Rand & L. S. Young), 366–381. Springer-Verlag, Berlin, Germany.

Theiler, J. (1986) Spurious dimensions from correlation algorithms applied to a limited time-series data. *Phys. Rev. Lett. A* **34**, 2427.

Theiler, J. (1991) Some comments on the correlation dimension of a 1/f-alpha noise. *Phys.Rev. Lett.A* **155**, 480.

Van Kampen, N. G. (1981) *Stochastic Processes in Physics and Chemistry*. North-Holland Physics Publishing, Amsterdam, The Netherlands.

Wiener, N. (1938) The homogeneous chaos. *Am. J. Math*, **60**, 887–936.