Proceedings WARMICE-2012
Sudhir Kumar, et al. (eds)
I.A.H., 2012

A Chaotic Analysis of Koyna Reservoir Catchment Rainfall Data

T.A. Fathima and V. Jothiprakash
Department of Civil Engineering, Indian Institute of Technology Bombay, Powai, Mumbai – 400076.

Abstract : The significance of treating rainfall as a chaotic system instead of a stochastic system is gaining interest in recent past studies and helps to understand the dynamic behaviour of the processes in a better way. Out of various chaotic methods available to analyse time series data, correlation dimension method is employed in the present study. Correlation dimension method is reported as a better method compared to other methods, especially when the data set is large and noisy. The scaling region in the Heaviside step function has to be identified to find the correlation exponent. The slope of the scaling region gives the correlation exponent value for various embedding dimension (m). Based upon the relationship, the behaviour of the system can be classified as chaotic (increasing and then saturating correlation exponent), stochastic (increasing correlation exponent with increase in embedding dimension) or deterministic (constant correlation exponent). The average daily rainfall observed at Koyna reservoir catchment area for a period of 49 years (1961-2009) in Maharashtra, India, has been taken up as a case study. From the detailed non-linear dynamic analysis using correlation dimension method it is found that the Koyna Reservoir catchment rainfall is showing a chaotic behaviour. The nearest whole number value of the saturation value of the correlation exponent is one, indicating that the correlation dimension for the daily average rainfall is one. This also indicates that the behaviour is low chaotic and high deterministic in nature.

Key words : Non linear chaotic analysis, chaotic behaviour, correlation dimension method, Koyna catchment average rainfall, correlation integral, correlation exponent, embedding dimension, correlation dimension.

INTRODUCTION

Lorenz (1963) was the first to point out possible chaotic characteristics of atmosphere and climate. A major breakthrough in chaos research came when several independent researchers proposed methods to use empirical time series to evaluate the dimension of the attractor in the system underlying such series (Packard et al,1981; Takens, 1981; Grassberger and Procaccia, 1983). The dimension of the attractor is a way to mathematically characterize the 'complexity' of the system's evolution. A series of investigations conducted subsequently on various climatic data reported the presence of low-dimensional attractors (Nicolis et al 1983; Tsonis and Elsner, 1988). These results, however, did not lack criticisms and responses (Lorenz,1991; Sivakumar,2002), since many of these earlier studies involved rather short and noisy time series and often used only one method to identify chaos. During the twentieth century many mathematicians and scientists contributed parts of today's chaos theory.

Chaos theory is, very generally, the study of forever changing complex systems (dynamic system) based on mathematical concepts of recursion. Chaos theory attempts to explain the fact that complex and unpredictable results can and will occur in systems that are sensitive to their initial conditions (Williams, 1997). Thus chaos deals with long-term evolution of how something changes over a long time. A chaotic time series looks irregular. The chaos that is well reported is a particular class of how something changes over time. In fact, change and time are the two fundamental subjects that together make up the foundation of chaos. Chaos theory attempts to explain the fact that complex and unpredictable results can and will occur in systems that are sensitive to their initial conditions (Alligood et al, 1997). The validity of the application of chaos to a data depends on many conditions, the length of the data being the most important of all (Sivakumar et al, 1998). Sivakumar et al (2000) discussed these issues with particular reference to hydrologic time series, while a review of chaos studies in the much broader spectrum of geophysics is presented by Sivakumar (2004). Some of the works on rainfall, runoff and river flow are depicted in the following Table 1. Most of these studies, however, have dealt with different observables from global, regional, or local climatic characteristics and studies on regional hydrology and chaos are almost non-existent. The study of chaos helps to identify the minimum number of dimensions sufficient to model the nonlinear dynamics of the reservoir rainfall.

BASIC TERMS AND PARAMETERS USED IN CHAOS

The basic terms which are used in Chaos are Phase Space, Lag, Embedding Dimension, Correlation integral, Correlation exponent, Correlation dimension and Attractor. These basic terms are explained as follows:

Phase Space

A standard phase space is the graph between different kinds of variables. But we are using pseudo phase space diagram for determining chaos. Pseudo phase space diagram is an imaginary graphical space in which the axes represent values of just one physical feature, taken at different times. The pseudo space co-ordinates represent the variables needed to specify the state of a dynamical system. Some time series can be very long and therefore difficult to show on a single graph, a phase space plot condenses all the data into a manageable space on a graph. Although no one can draw a graph for more than three variables while still keeping the axes at right angles to one another, the idea of phase space holds for any number of variables.

Lag

Pseudo phase space is for comparing a time series to later measurements within the same data (a subseries). For instance, a plot of x_{t+1} versus x_t shows how each observation (x_t) compares to the next one (x_{t+1}). In that comparison, the group of x_t values is called as the basic series and the group of x_{t+1} value as the subseries. By extending that idea, one can also compare each observation to the one made two measurements later (x_{t+2} versus x_t), three measurements later (x_{t+3} versus x_t), and so on. The displacement or amount of offset, in units of number of events, is called the lag. The displacement or amount of offset, in units of number of events, is called the lag. Lag is selected as a constant interval in time between the basic time series and any sub series we're comparing to it. For instance, the sub series $x_{t+1}$is based on a lag of one, x_{t+2} is based on a lag of two, and soon. Lagged phase space is a special type of pseudo phase space in which the coordinates represent lagged values of one physical feature.

Embedding Dimension

Pseudo (lagged) phase spaces graph can have two or three axes or dimensions. Chaologists often extend the idea of phase space to more than three dimensions. In fact, one can analyze mathematically and compare any number of subseries of the basic data. The embedding dimension is the total number of separate time series (including the original series, plus the shorter series obtained by lagging that series) included in any one analysis.

Table 1. Chaos studies in rainfall runoff and river flow

Data	Correlation dimension	Reference
Monthly rainfall in Nauru Island	2.5-4.5	Hense (1987)
15-s rainfall intensity in Boston, MA, USA	3.78	Rodriguez-Iturbe et al(1989)
Weekly rainfall in Genoa	No low dimension	Rodriguez-Iturbe et al (1989)
15-s rainfall intensity in Boston (3 stations)	3.35, 3.75, 3.60	Sharifi et al (1990)
10-s rainfall intensity generated from cloud model	1.5	Islam et al (1993)
Daily rainfall in Hong Kong (three stations)	0.95, 1.76, 1.65	Jayawardena and Lai (1994)
15-min rainfall in Greece		
	No low dimension	Koutsoyiannis and Pachakis (1996)
Daily rainfall in Singapore (six stations)		Sivakumar et al (1998),
	1.01, 1.03, 1.06,	Sivakumar (1999)
Monthly rainfall in Gota River, Sweden	1.03,1.02, 1.06	Sivakumar et al (2000) Sivakumar
Monthly runoff coefficient in Gota River,	6.4	et al (2000)
Sweden	7.8	
Monthly flow in Gota River, Sweden		Sivakumar et al (2000) Sivakumar
Daily flow in Mississippi River, MO, USA	2.32	(2001)
	5.5	

Correlation Integral (Cr)

It is the ratio of number of pairs coming inside a circle of radius r (any high frequency value of rainfall) with respect to each and every co ordinate in phase space and the total number of pairs in the phase space. It gives an idea of correlation of one pair to another pair in phase space. Cr will go on increasing with increase in the radius.

Correlation exponent (í)

After finding the Cr, plot log r Vs log Cr for various embedding dimension and get the slope of the most straight line portion of the graph. This will give correlation exponent for each embedding dimension. These slope values are employed to find the correlation dimension of the chaotic time series.

Correlation dimension

Plotting the correlation exponent values against each embedding dimension will give a saturation value of correlation exponent after a particular value of embedding dimension. After that the correlation exponent remains same even after with increasing embedding dimension when the time series is having a chaotic nature. This saturating value of correlation exponent is termed as correlation dimension which is the dimension required for modeling the entire system or the future prediction.

Attractor

It is a geometric object which incorporates the properties of the entire system in the long run in the phase space. It characterizes the long term behaviour of a system in the phase space. If the dynamics of the system can be reduced to a set of deterministic laws, the trajectories of the system converge towards the subset of the phase space, called the attractor (Henon, 1976).

STUDY AREA AND TIME SERIES

The Koyna catchment area average daily rainfall observed for a period of 49 years (1961-2009) has been used to study its behaviour. The Koyna reservoir is situated at the west coast of Maharashtra, India and lies between north latitude of $17^0$0' to $17^0$59' and east longitude of $73^0$02' to $73^0$35'. The location of the study area of Koyna reservoir is shown in Figure 1. Figure 2 depicts the time series plot of 49 years of daily observed rainfall. From the time series plot, it is observed that there is rainfall only during monsoon period, from June to October every year indicating that the reservoir is located in an intermittent river system. Thus from the Figure 2, it can be seen that seven months in a year the rainfall is zero posing a great difficulty to any type of rainfall prediction models. It is found that the full year rainfall data is having 61% of zero values. However within the monsoon months it is only 8% zero values which are scattered in five monsoon months.

METHODOLOGY

Out of various methods employed to study the chaotic behaviour of the time series, correlation dimension method is one of the most widely used method where data length is large and can have some noise in it. The procedure usually begins

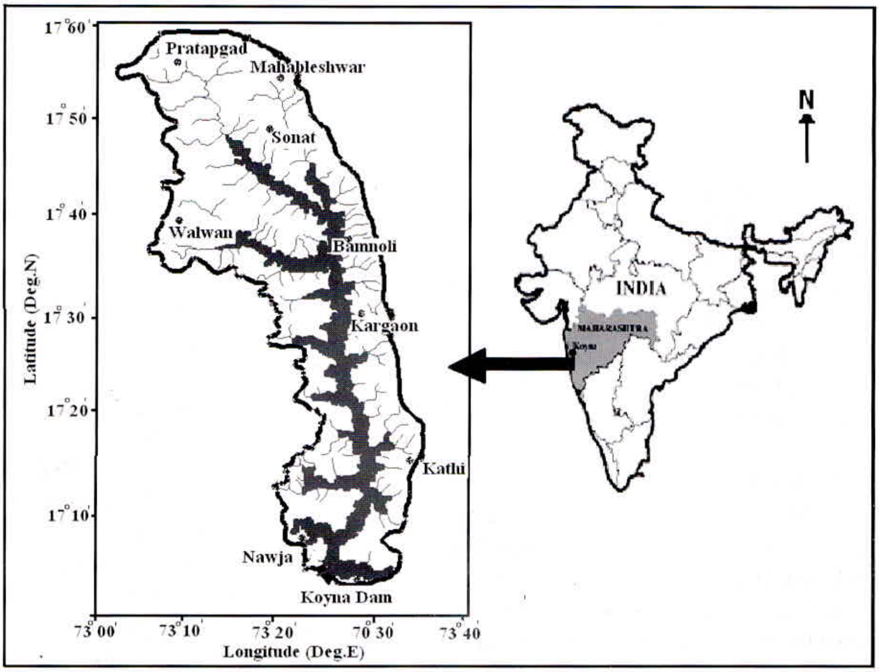

Fig. 1. Location of Koyna reservoir

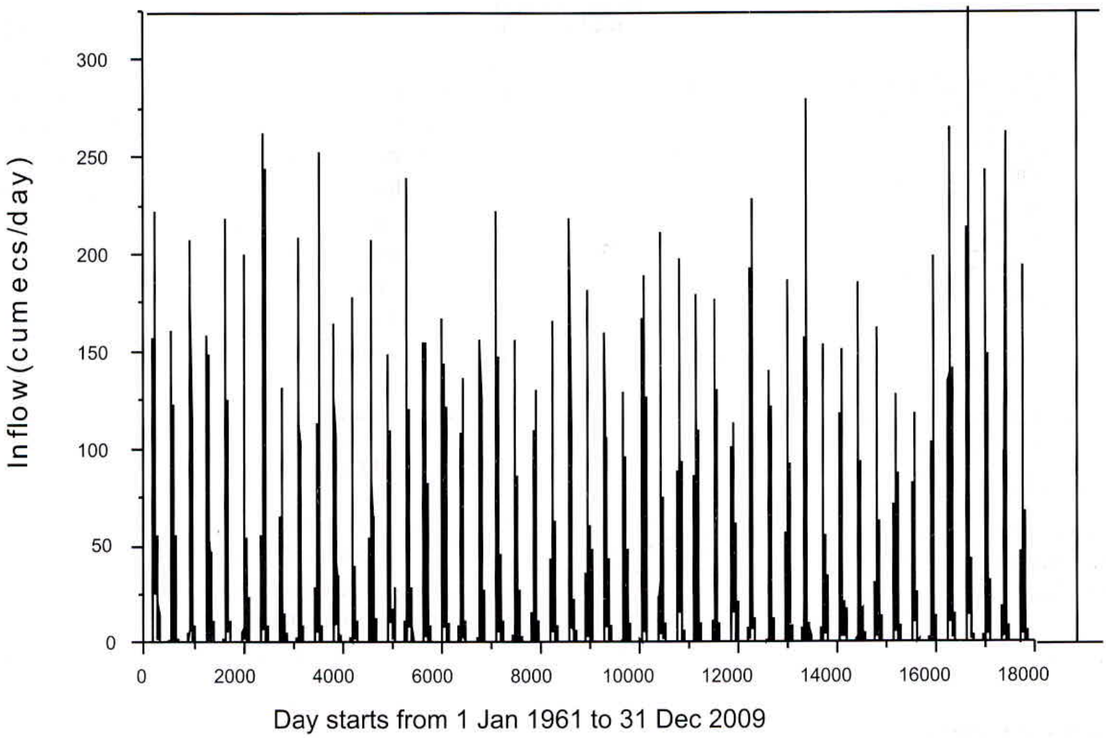

Fig. 2. Time series plot of Full year Rainfall

by embedding the data in a two-dimensional pseudo phase space. For a given radius r, count the number of points within distance r from the reference point. After doing that for each point on the trajectory, sum the counts and normalize the sum. That yields a so-called correlation sum. Then repeat that procedure to get correlation sums for larger and larger values of r. A log plot of correlation sum versus r (for that particular embedding dimension) typically shows a straight or nearly straight central region. The slope of that straight segment is the correlation dimension. The next step is to repeat the entire procedure for larger and larger embedding dimensions. For chaotic data, the correlation dimension initially increases with embedding dimension, but eventually (at least in the ideal case) it asymptotically approaches a final (true) value.

Step by step procedure for finding correlation dimension (Grassbereger-Procassia, 1983)

1. Choose the time delay which is equal to one.
2. Select the given inflow data as basic series Y_t and find out the subseries with different lags say lag 1 to lag 20(in the present study). The dynamics of time series of inflow $\{x_1,x_2,....x_n\}$ are fully captured or embedded in the m-dimensional phase space (assuming delay time as 1) defined by equation 1

$$Y_t=\{x_t,x_{t-1,}x_{t-2,}...x_{t-(m-1)}\} \quad (1)$$

3. Draw the phase space diagram from one dimension to 20 dimensions. Consider each phase space diagram with different dimension individually and fix a radius r.The

steps in radius used in the present study are 0.05,1,2,3 ,4 and 5.Fix the dimension m(say m=1). Fix any particular radius step (say 0.05) and find the number of points coming inside the circle. Increase the radius by constant interval (steps in radius) and find the number of points coming inside the circle. The radius should be increased to a particular level in constant interval such a way that all points in the phase space diagram should come under that circle. Then with the same m try for different steps of r as mentioned above. Then try with the next m value and continue the procedure till m=20.

Find C(r) values for different values of r by increasing steps in constant intervals by using the following equation 2.

$$C(r) = \lim_{n\to\infty} \frac{2}{N(N-1)} \sum_{i,j(1<i<j<N)} H(r - |Yi - Yj|) \qquad (2)$$

where H is the Heaviside step function, with H(u) =1 for u >0, and H(u)= 0 for $u \leq 0$,

N= Number of data points; u = $r - |Yi - Yj|$;

r is the radius of the sphere centered on Y_i or Y_j. Y_i is the basic series and Y_j is the subseries of inflow data

4. Plot correlation sum versus radius on log paper, defining a separate relation for each embedding dimension.
5. Measure the correlation dimension as the slope of the (hopefully) straight middle zone of log C(r) Vs log r graph (the scaling region) of each relation. The slope called correlation exponent (õ) is found from equation 3.

   $$õ = \log C(r) / \log r \qquad (3)$$

6. Plot correlation dimension versus embedding dimension on arithmetic paper.
7. Decide if the latter relation seems to become asymptotic (hence possibly indicating an attractor) or keeps increasing in proportion to the embedding dimension (hence probably random data). The next nearest whole number value of saturation value of õ gives the correlation dimension for the chaos. This correlation dimension value gives the number of variables required to predict the future value (Williams 1997)
8. The correlation dimension is also found out by local slope method by plotting log r Vs correlation exponent(õ) .The value at which plateau forms gives the correlation dimension(Sivakumar, 2000).

RESULTS AND DISCUSSION

The above discussed correlation dimension method has been applied for Koyna catchment average rainfall series. The results obtained from this analysis are explained below.

Reconstruction of Phase space plot

It is worth mentioning that phase space diagram also gives an idea about the attractor. The attractor indicates the possibility of chaos (chaotic behaviour) in the data set. The phase space plot of full year catchment rainfall is shown in Figure 3 from which one can find the attractor region.

Estimation of correlation Integral

The correlation integral is determined by the Equation 3 after the reconstruction of phase space with different dimensions. By finding the logarithm of r as well as Cr, plot the graph between log r Vs log Cr for various dimensions. The log r Vs log Cr plot with a radius equal to three is shown in Figure 4 . The value of Cr decreases from 0.47 for m=1 to 0.29 for m=20.But it increases up to 1 with increase in radius.

Estimation of correlation dimension by plotting m Vs slope

Correlation exponent is the slope of scaling region of log r Vs log Cr graph. The scaling region is the straight line portion of the log r Vs log Cr graph. By fitting straight line to the approximate linear

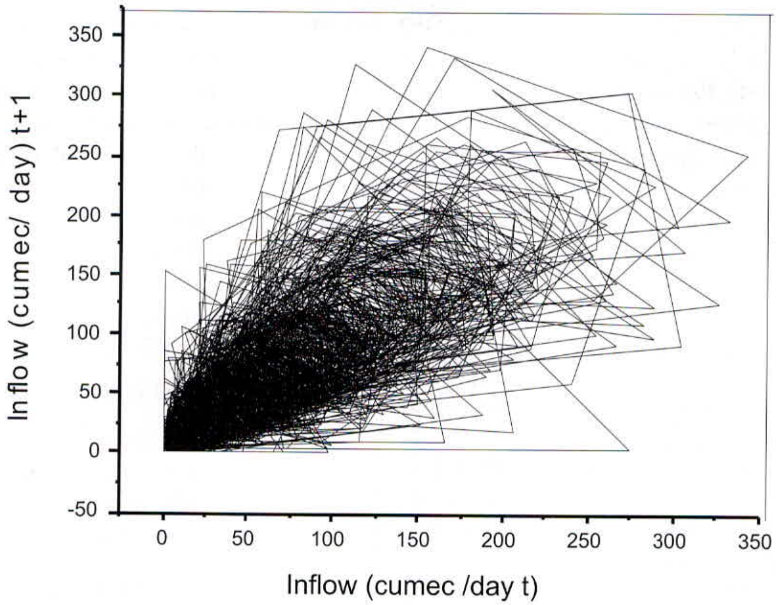

Fig. 3. Phase space plot of daily full year reservoir rainfall

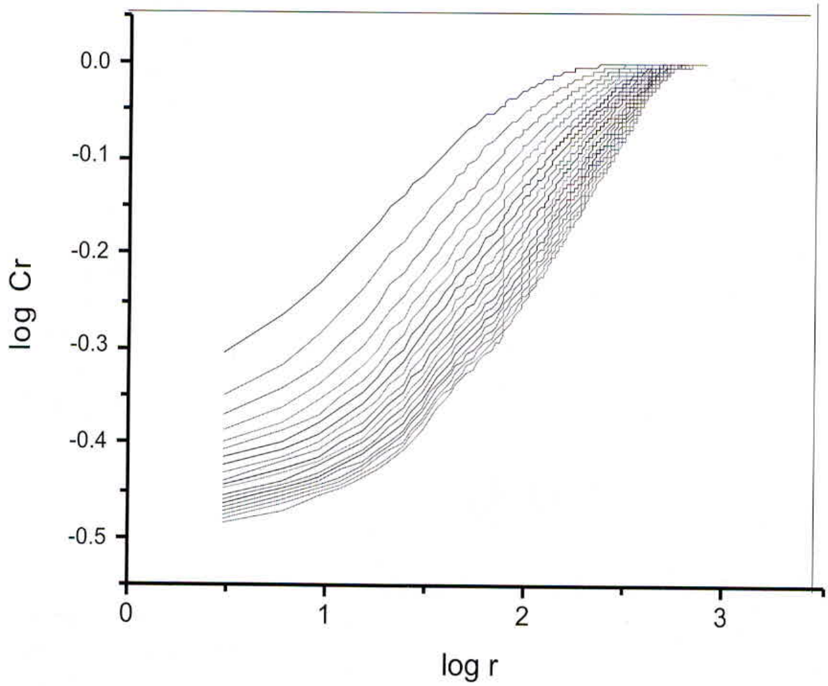

Fig. 4. log r versus log Cr graph

portion of log r Vs log Cr graph, the slopes are obtained. Figure 5 is derived by this method. The correlation exponent resulted from the chaotic analysis of Koyna reservoir rainfall is shown in Figure 5.The embedded dimension worked out is up to 20. It is clear from the Figure 5 that there is low chaotic behaviour in the time series with a correlation dimension of 1. Hence the correlation dimension for daily full year rainfall data is 1. It shows that daily full year rainfall values are highly deterministic with low chaotic behaviour.

CONCLUSIONS

The present study aimed at categorising the Koyna catchment area average rainfall series into a stochastic, deterministic or chaotic series. Correlation dimension method has been used to analyse the chaotic behaviour. Forty nine years of daily historical average rainfall data from 1961-2009 pertaining to Koyna catchment area has been used in the present study. The 'r' value taken is in between the minimum observed rainfall other than zero and one third of the highest frequency rainfall values to get a better log r Vs log Cr graph for estimating the correlation exponent. An appropriate section has been selected in between the initial and converged points of log r Vs log Cr graph to estimate the correlation exponent. It is found that the straight line portion which is parallel to adjacent 'm' values will be the appropriate scaling region to be selected to estimate the slope. The correlation dimension of daily average rainfall series is found as 1 which indicates that only one dimension is required to model the system. Thus it may be concluded that the Koyna area catchment rainfall is highly deterministic with low chaotic behaviour.

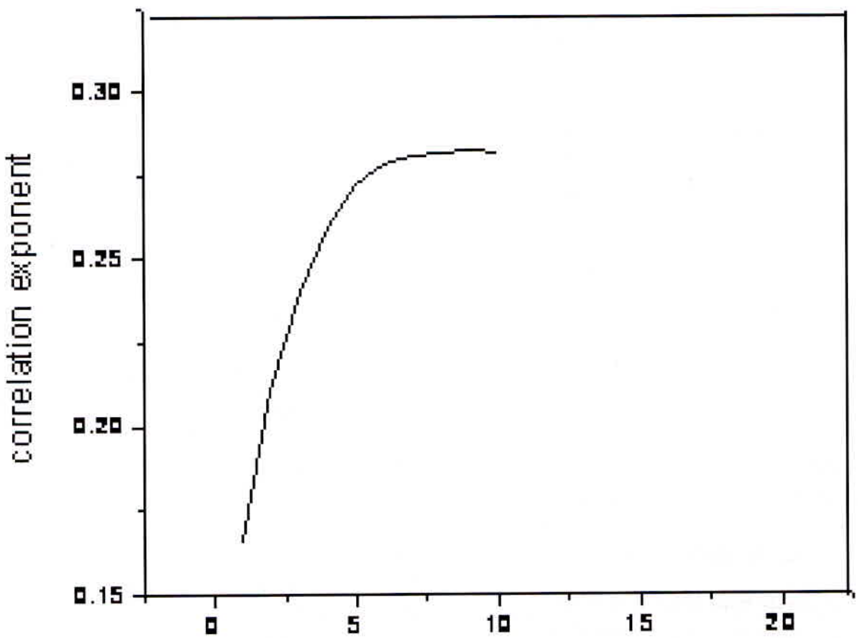

Fig. 5. Embedding dimension Vs correlation exponent plot

REFERENCES

Alligood,K.,Saucer,T. and Yorke,J.A.(1997). *An Introduction to Dynamical Systems.* Springer-Verlag,Newyork.

Grassberger,P. and Procaccia,I.(1983). Measuring the strangeness of strange attractors. *Physica D*,Vol.9;pp.189–208.

Hénon, M. (1976). A two-dimensional mapping with a strange attractor. *Communications in Mathematical Physics* , Vol. 50;pp.69-77.

Hense, A.(1987), On the possible existence of a strange attractor for the southern oscillation. *Beitr Phys Atmos,* Vol.60,No:1;pp.34–47.

Islam, S., Bras, R. L. and Rodriguez-Iturbe, I.(1993). A possible explanation for low correlation dimension estimates for the atmosphere. *J Appl Meteor* , Vol. 32;pp.203–208.

Jayawardena, A.W. and Lai,F.(1994). Analysis and prediction of chaos in rainfall and stream flow time series. *Journal of Hydrology*,Vol 153;pp.23-52

Koutsoyiannis,D. and Pachakis,D.(1996). Deterministic chaos versus stochasticity in analysis and modeling of point rainfall series. *JGeophys Res*, Vol.10, No:D21;pp.51-26441.

Lorenz,E.N.(1991). Dimension of weather and climate attractors. *Nature,*Vol.253, Sept,pp.3-241.

Lorenz, E. N.(1963). Deterministic non periodic flow, *Journal of the Atmospheric Sciences.*, Vol.20, No:2;pp.130-141.

Nicolis, J., Mayer-Kress, G. and Haubs, G. (1983), Non-Uniform Chaotic Dynamics with Implications to Information Processing, *Zeitschrift für Naturforschung.*, Vol.38a;pp.69-1157.

Packard,N.H .,Crutchfiled, J.P.,Farmer, J.D. and Shaw,R.S. (1981). Geometry from a time series. *Physical Review Letters*, Vol.45' No:9;pp.712-716.

Rodriguez-Iturbe ,I., De Power, F.B., Sharifi,M.B. and Georgakakos,K.P. (1989). Chaos in rainfall.*Water Resour Res*,Vol.25,No:7;pp.167-1667

Sharifi, M.B., Georgakakos, K.P. and Rodriguez-Iturbe,I. (1990). Evidence of deterministic chaos in the pulse of storm rainfall. *J Atmos Sci* ;Vol.47;pp.888-893.

Sivakumar, B.,Liong, S.Y. and Liaw, C.Y.(1998). Evidence of chaotic behavior in Singapore rainfall. *J Am Water Resour Assoc*,Vol.34,No:2;pp.10-301.

Sivakumar,B.(1999). Chaos theory in hydrology: Important issues and interpretations. *Journal of Hydrology*, Vol.227;pp.1-20

Sivakumar, B., Berndtsson, R., Olsson, J., Jinno, K. and Kawamura, A. (2000). Dynamics of monthly rainfall-runoff process at the Gota basin: a search for chaos, *Hydrol Earth Syst Sci*,Vol.4,No:3;pp. 17-407.

Sivakumar,B.(2001). Rainfall Dynamics at Different Temporal Scales:A Chaotic Perspective, *Hydrology and Earth System Sciences*,Vol.5,No:4;pp.645-651

Sivakumar, B. (2004). Chaos theory in geophysics: past, present and future Chaos. *Solitons and Fractals*,Vol.19;pp.441–462

Sivakumar ,B. (2002). Discussion on "Analysis of cross-correlated chaotic streamflows" by Elshorbagy et al. *Hydrol Sci J*;Vol.47,No:3;pp.7-523.

Takens, F.(1981). *Detecting strange attractors in turbulence,* Vol.898,pp.81-366, Berlin, Springer.

Tsonis,A.A. and Elsner,J.B.(1988). The weather attractor over very short timescales. *Nature*,Vol.333;pp.7-545.

Williams,P(1997). *Chaos theory Tamed*, US Geological Survey (Ret.),Joseph Henry Press.,Washington.