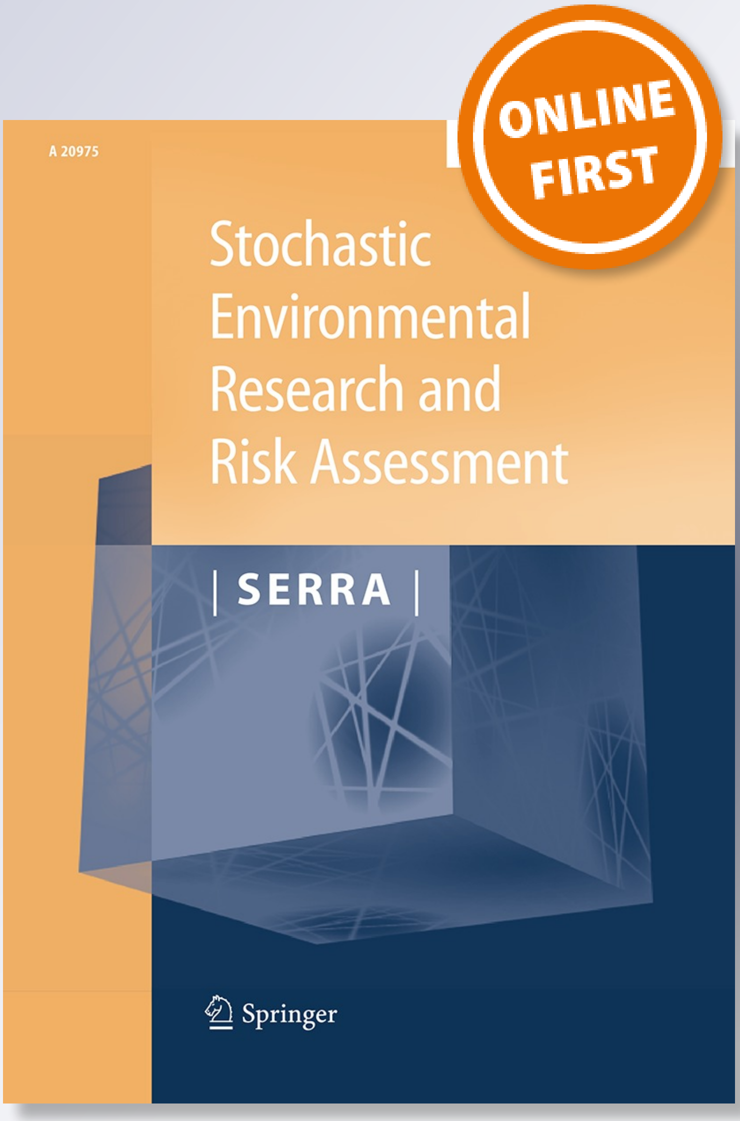
Nonlinear analysis of rainfall variability in Australia

Bellie Sivakumar, Fitsum M. Woldemeskel & Carlos E. Puente

Stochastic Environmental Research and Risk Assessment

ISSN 1436-3240

Stoch Environ Res Risk Assess
DOI 10.1007/s00477-013-0689-y



Stoch Environ Res Risk Assess
DOI 10.1007/s00477-013-0689-y

ORIGINAL PAPER

Nonlinear analysis of rainfall variability in Australia

**Bellie Sivakumar · Fitsum M. Woldemeskel ·
Carlos E. Puente**

Abstract This study examines the utility of nonlinear dynamic concepts for analysis of rainfall variability across Western Australia. The correlation dimension method, which involves data reconstruction and nearest neighbor concepts, is employed to estimate the dimensionality of rainfall time series and assess the degree of variability (or complexity) of rainfall dynamics. Monthly rainfall data observed over a period of 67 years (January 1937–December 2003) from each of 62 raingage stations across Western Australia are analyzed. The results indicate the utility of the dimensionality concept for identification of rainfall variability at point locations and rainfall patterns across space. The dimension estimates (ranging from 4.63 to 8.29) suggest a medium-level complexity in rainfall dynamics at each of the 62 stations and, hence, across Western Australia. Rainfall dynamics are generally more complex in the southwest part of Western Australia than in other parts of the state, with the far south exhibiting a greater complexity. Also, the mid-western region exhibits slightly more complex rainfall dynamics when compared to the northeast. The reliability of the correlation dimension estimates with respect to data size (i.e. number of data points in time series) is also discussed. The outcomes of the present study, the first ever based on nonlinear dynamic and chaos theories for rainfall in Australia and also across a far more extensive network of raingages, can provide new avenues for rainfall (and other hydrologic) studies in Australia and for spatial/spatio-temporal analysis.

Keywords Rainfall variability · Nonlinear dynamics · Phase space reconstruction · Correlation dimension · Australia

1 Introduction

Australia is the driest inhabited continent in the world, with a mean annual rainfall of only about 465 mm. As a consequence, water is a critical resource, and its proper planning and management is vital for sustaining the nation's people, ecosystems, environment, and economy. However, the extreme variability of rainfall in both space and time brings tremendous challenges to this, considering also the significant variability in the distribution of population across the country (e.g. very high density in a few big cities in the coastal regions vs. very low density in the largely rural towns in the interior regions). For instance: (1) coastal regions generally receive far greater rainfalls than interior regions; (2) northeast Australia generally gets significantly higher rainfalls than the rest of the country, especially central and western parts; and (3) much of the rainfall occurs during the three summer months (December–February).

Under these circumstances, sustainable planning and management of water is one of the biggest challenges facing the nation, especially regions that receive very low rainfall, such as Western Australia. This problem is even further complicated by two very likely future scenarios: (1) population growth; and (2) climate change impacts. Some current estimates reveal that the nation's population will increase from about 22.7 million at present to about 36 million by around 2050 (Australian Bureau of Statistics

---

B. Sivakumar · F. M. Woldemeskel
School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia
e-mail: s.bellie@unsw.edu.au

B. Sivakumar (✉) · C. E. Puente
Department of Land, Air and Water Resources,
University of California, Davis, CA 95616, USA
e-mail: sbellie@ucdavis.edu

2008), if the current trends in fertility, life expectancy at birth, net overseas migration, and other relevant factors continue. This growth in population suggests not only increases in water demands but also significant changes to land use and development. The population estimates also reveal that big cities will again bear the brunt of this growth (including Perth, the capital of Western Australia). Future climate change projections generally indicate increased temperatures, reduced rainfall, increased frequency and intensity of extreme hydroclimatic events (e.g. floods, droughts, sea level rises), and negative effects on water quality in Australia (CSIRO and Bureau of Meteorology 2007; Chiew et al. 2010; Cleugh et al. 2011). Since rainfall plays a central role in all water resources planning and management activities as well as in many others, adequate knowledge of its dynamics (spatial and temporal) is crucial for preparedness, mitigation, and adaptation efforts.

Numerous studies have investigated the variability of rainfall in Australia (e.g. Lavery et al. 1997; Jeffrey et al. 2001; Beesley et al. 2009; Jones et al. 2009; Mehrotra and Sharma 2009, 2010; Taschetto and England 2009; Western et al. 2011). The studies have addressed various aspects of rainfall, including statistics, interpolation, scaling, and, more recently, in terms of climate change impacts. They have certainly provided encouraging outcomes, especially in terms of their efforts to obtain better quality rainfall data and related hydroclimatic information for more detailed studies. However, they have not given sufficient attention to the nonlinear, and especially chaotic, nature of rainfall in Australia. Since, generally speaking, rainfall dynamics (and climate dynamics at large) are inherently nonlinear in nature and may possibly exhibit chaotic behavior (e.g. Lorenz 1963; Tsonis and Elsner 1988; Rodriguez-Iturbe et al. 1989; Puente and Obregon 1996; Sivakumar et al. 2001; see also Sivakumar 2000, 2004, 2009 for reviews), study of the nonlinear and chaotic aspects may offer additional insights into rainfall dynamics in Australia and help improve water resources planning and management. Furthermore, although many studies have employed the nonlinear dynamic concepts for studying rainfall dynamics in many parts around the world, including the United States (e.g. Rodriguez-Iturbe et al. 1989; Puente and Obregon 1996; Sivakumar et al. 2001, 2006), Sweden (e.g. Berndtsson et al. 1994; Sivakumar et al. 2000), Hong Kong (e.g. Jayawardena and Lai 1994), Singapore (Sivakumar et al. 1999a), Thailand (e.g. Jayawardena and Gurung 2000), Greece (e.g. Koutsoyiannis 2006), and Korea (e.g. Kyoung et al. 2011), such studies have largely been limited to rainfall data observed at only one or a few stations at the most. However, since water resources planning and management at state and national levels requires knowledge of rainfall dynamics over a much greater spatial extent, it is crucial to study rainfall data observed at a large number of stations in the area.

In view of these, the present study examines the utility of nonlinear dynamic concepts for analysis of rainfall variability in Australia, and more specifically Western Australia. Since monthly scale is arguably the most appropriate scale for medium- to long-term water resources studies, especially those related to water shortages and droughts, monthly rainfall dynamics are studied. Rainfall data observed over a period of 67 years (January 1937–December 2003) from each of 62 raingage stations across Western Australia are considered. These data are analyzed using correlation dimension method, a nonlinear dynamic method having its base on data embedding and nearest neighbor search procedures. The main purpose herein is to determine the dimensionality of the rainfall dynamics at each of the 62 stations (represented through the variability of the time series of observed rainfall data) so as to assess the variability of rainfall across Western Australia. The dimensionality estimation provides information on the number of variables dominantly influencing the rainfall dynamics and, consequently, helps in the identification of appropriate model complexity.

The rest of this paper is organized as follows. Section 2 presents a brief description of the geography and climate of Western Australia and describes the monthly rainfall data from 62 stations considered in this study. Section 3 details the correlation dimension method. The results from the correlation dimension analysis of rainfall data are presented and discussed in Sect. 4. Conclusions and future research directions are reported in Sect. 5.

2 Study area and data

2.1 Regional information

The climate and, hence, rainfall in Australia are greatly influenced by the surrounding oceans. For example, the El-Niño Southern Oscillation (ENSO), the western Pacific and Indian Ocean sea surface temperatures, and the Southern Ocean atmospheric variability influence the different regions of Australia by varying degrees (Taschetto and England 2009). The climate varies from tropical in the north to arid in the middle to temperate in the south and, consequently, rainfall is highly variable in both space and time. A large part of the country is dry, especially the middle and the west. More than 80 % of the country gets an annual rainfall of less than 600 mm, but the tropical region of the far north receives an annual rainfall of over 4,000 mm.

Western Australia (Fig. 1) is Australia's largest state by areal extent, with an area of 2,645,615 km^2 (land area 2,529,875 km^2). It is bordered by the Indian Ocean to the north and west, the Great Australian Bight and the Indian Ocean to the south, the Northern Territory to the northeast, and South Australia to the southeast. The current population of Western Australia is about 2.3 million, which is about 10 % of the nation's total population; however, the state's population is projected to double over the next four decades, reaching about 4.3 million people by 2050. About 85 % of the current population (i.e. about 1.95 million) live in the southwest corner of the state, with the capital city Perth alone having 1.7 million.

As for climate in the state, the southwest coastal area has a Mediterranean climate. The average annual rainfall here varies from 300 mm at the edge of the Wheatbelt region (one of the nine regions in the state) to 1,400 mm in the wettest north areas. However, in the months of November to March, evaporation exceeds rainfall, and it is generally very dry. A major reduction in winter rainfall has been observed since the mid-1970s, with a greater number of extreme rainfall events in the summer months (e.g. Cleugh et al. 2011). The central four-fifth of the state is semi-arid or desert and is lightly inhabited with the only significant activity being mining. Here, the annual rainfall averages 200–250 mm, most of which occurs in sporadic torrential falls related to cyclone events in summer months, northern tropical regions excepted. The Kimberly region (in the north) has an extremely hot monsoonal climate with average annual rainfall ranging from 500 to 1500 mm, but there is a very long almost rainless season from April to November.

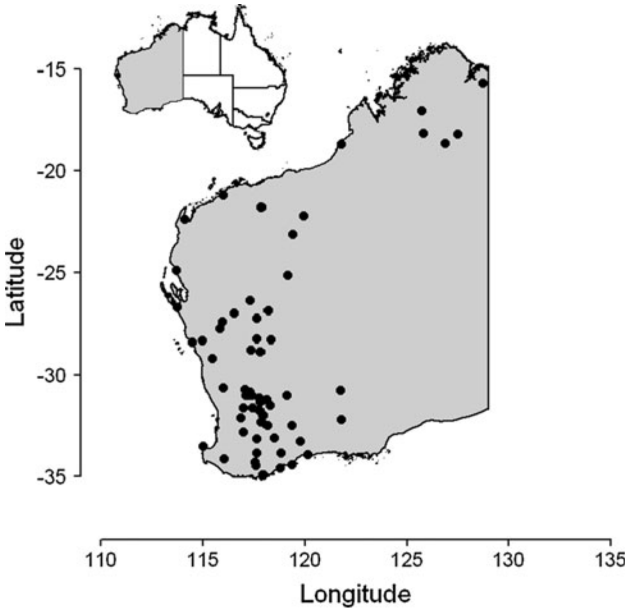


Fig. 1 Map of Australia and locations of 62 raingage stations in Western Australia

2.2 Rainfall data

In Australia, rainfall is mainly monitored using the thousands of raingages installed at various locations across the country, by the Australian Bureau of Meteorology. However, the majority of these raingages are only near the coasts of northeast and east (including the cities of Cairns and Brisbane), southeast (including the cities of Sydney and Melbourne) and southwest (including Perth) of the country. These regions collectively are where the population is heavily concentrated and also have most of the big cities. In the interior of the country, which is largely desert and very sparsely populated, the number of raingages is very few.

The rainfall archive of the Australian Bureau of Meteorology contains more than 17,000 raingage stations (Oke et al. 2009). The length and period of rainfall measurements in these stations differ considerably. For instance: (1) some stations have records for more than 100 years, while some others have records for much shorter periods; and (2) some stations have started recording only recently, while recording has already ended in some others.

In the present study, monthly rainfall observed over a period of 67 years (January 1937–December 2003) at each of 62 raingages across Western Australia are considered (see Fig. 1; Table 1). These 62 stations are selected from high-quality monthly rainfall measuring raingages, identified by Lavery et al. (1997). There are a few missing rainfall measurements in some stations (see footnote of Table 1 for missing stations and months). These missing data are filled by the mean rainfall of the respective month (calculated from all the available years) at the respective station. Table 1 presents some basic information about these stations (station number, latitude, longitude, and elevation) and statistics of monthly rainfall values (mean, standard deviation, maximum, and percentage of zeros).

The information in Table 1 offers some useful insights on station characteristics (especially topography) and monthly rainfall statistics. As for station characteristics, the elevation ranges from as low as 4 m (Station 6011) to as high as 670 m (Station 7169). In terms of monthly rainfall statistics, the mean ranges from 16.90 mm (Station 6048) to 73.53 mm (Station 9619), standard deviation from 24.32 mm (Station 6048) to 100.06 mm (Station 3017), maximum from 131.80 mm (Station 10062) to 764.30 mm (Station 3030), and no-rainfall months from 0 % (Station 9564) to 40.92 % (i.e. 329 months) (Station 5014). These statistics clearly reveal the significant spatial variability of rainfall in Western Australia, including indicating a four-to-six-fold difference in the range (maximum and minimum) of values in mean, standard deviation, and maximum rainfall observed among all the 62 stations, and an even greater difference in terms of number of zeros. For a better

Table 1 Raingage station details and basic statistics and correlation dimension estimates of rainfall data from Western Australia (Period: January 1937–December 2003)

S. No.	Station No.	Lat.[b]	Long.[c]	Elev. (m)[d]	Mean (mm)[e]	Std. Dev. (mm)[f]	Max. (mm)[g]	% Zeros[h]	Cor. Dim.[i]
1	2014	−15.65	128.71	31	67.63	98.14	575.9	33.71	5.44
2	2019	−18.63	126.86	280	39.01	68.43	494.7	38.06	4.63
3	2020[a]	−18.19	127.50	430	45.45	78.49	503.7	35.07	5.31
4	3017[a]	−17.05	125.70	270	64.31	100.06	686.8	36.19	4.80
5	3027	−18.14	125.78	120	44.07	78.81	513.9	39.55	4.79
6	3030	−18.68	121.78	11	42.61	88.33	764.3	36.57	4.87
7	4006[a]	−22.18	119.93	480	27.57	52.08	649.6	36.32	6.17
8	5004	−22.38	114.11	15	22.62	44.56	402.8	36.69	4.96
9	5008[a]	−21.19	115.98	11	25.68	53.86	675.2	40.42	5.19
10	5014	−21.79	117.86	450	32.22	59.14	449.0	40.92	5.70
11	6011	−24.88	113.67	4	18.36	29.82	206.0	20.40	5.17
12	6045	−26.70	113.71	50	23.98	35.96	252.4	33.83	5.62
13	6048	−27.39	115.97	260	16.90	24.32	178.6	25.62	6.59
14	6055	−27.75	115.83	300	19.50	24.95	152.6	19.65	6.97
15	7007	−26.98	116.54	300	17.87	25.89	184.1	19.78	6.72
16	7011	−28.79	117.36	400	20.00	26.53	249.8	17.79	7.08
17	7014[a]	−27.26	117.65	500	19.25	29.76	241.4	21.52	6.09
18	7020	−26.88	118.17	400	17.38	27.59	209.7	18.66	4.83
19	7049[a]	−26.37	117.33	400	17.00	28.53	350.5	29.35	5.09
20	7080[a]	−25.13	119.15	520	19.21	31.43	267.8	29.98	5.33
21	7090[a]	−28.84	117.83	450	22.72	29.21	266.7	20.15	7.26
22	7095	−28.23	117.65	400	20.08	25.98	202.7	16.79	8.29
23	7169[a]	−23.10	119.37	670	25.89	45.44	292.7	28.61	5.32
24	7197[a]	−28.28	118.31	400	17.47	25.59	216.0	25.50	6.92
25	8066	−30.70	117.06	310	27.72	28.43	190.9	9.33	6.99
26	8088	−29.19	115.44	153	33.05	38.20	271.3	10.07	6.44
27	8091	−30.64	116.01	203	37.68	38.64	244.9	5.97	6.71
28	8141[a]	−28.38	114.45	100	36.82	44.47	267.4	14.68	5.93
29	8147[a]	−28.33	114.96	270	28.65	33.16	206.2	13.68	6.44
0	9519[a]	−33.54	115.02	109	67.86	65.52	338.0	1.37	6.30
31	9520	−34.61	118.75	40	47.53	38.05	243.0	1.00	7.29
32	9557	−33.95	120.12	15	42.25	34.38	329.9	1.12	7.39
3	9561	−34.49	117.63	262	47.10	36.20	207.3	0.62	7.93
34	9564	−34.94	117.92	12	69.87	49.39	329.7	0.00	6.98
35	9594	−34.44	119.36	60	55.18	43.57	263.4	0.25	7.90
36	9619	−34.15	116.02	240	73.53	65.53	365.9	1.74	7.28
37	10032	−31.00	117.39	300	25.04	25.02	131.8	6.84	7.06
38	10037[a]	−31.73	117.76	300	25.36	25.49	147.8	10.57	7.51
39	10039	−31.01	117.20	295	25.79	25.99	153.7	9.33	7.34
40	10041[a]	−31.62	117.44	250	27.27	27.86	181.5	10.57	7.07
41	10045	−31.01	117.13	290	25.60	26.88	152.0	10.82	6.77
42	10091[a]	−31.63	117.01	195	29.50	30.61	185.6	11.32	6.78
43	10092	−31.48	118.28	315	27.20	25.80	143.7	6.09	7.06
44	10112	−31.19	118.10	292	25.65	25.22	158.2	7.71	7.13
45	10123	−31.99	117.93	250	25.35	25.78	182.1	11.19	6.96
46	10126	−31.12	117.79	290	26.79	26.81	157.5	6.22	7.48
47	10133	−30.83	117.32	320	24.26	24.89	137.6	7.46	7.05

Table 1 continued

S. No.	Station No.	Lat.[b]	Long.[c]	Elev. (m)[d]	Mean (mm)[e]	Std. Dev. (mm)[f]	Max. (mm)[g]	% Zeros[h]	Cor. Dim.[i]
48	10149	−31.34	117.82	320	26.18	26.31	167.2	6.84	7.07
49	10525	−33.85	117.64	328	36.36	31.65	236.0	3.73	7.57
50	10536	−32.33	117.87	295	30.43	29.96	206.7	6.09	7.40
51	10537[a]	−34.30	117.57	255	41.55	33.05	190.2	1.99	7.37
52	10541[a]	−33.86	118.81	280	31.40	27.70	232.7	3.73	6.92
53	10592	−33.10	118.46	286	28.85	27.58	214.3	3.23	6.51
54	10611[a]	−33.28	119.77	310	29.70	25.60	249.4	4.35	6.89
55	10636[a]	−33.16	117.67	340	33.01	31.54	251.4	6.72	6.89
56	10658[a]	−32.83	116.99	300	41.54	43.22	330.3	9.95	6.65
57	10668[a]	−32.47	118.14	280	27.00	27.44	211.8	10.95	7.40
58	10670	−32.50	119.36	310	28.32	26.27	204.4	6.59	6.93
59	10795[a]	−32.12	116.87	200	33.19	34.05	223.3	8.58	6.30
60	12011	−30.99	119.11	400	25.59	26.56	204.2	7.46	6.57
61	12013	−30.75	121.75	380	21.74	29.37	315.7	8.71	5.89
62	12065	−32.20	121.78	277	24.95	25.32	202.6	3.73	7.20

[a] One or a few data are missing at these stations. The month(s) of missing data is as follows. The missing data is obtained by averaging the corresponding monthly values during the other years of the records: Station 2020—Oct 01; Station 3017—Apr 96, Apr 98; Station 4006—Jul 98, Dec 98, Apr 03; Station 5008—Nov 03; Station 7014—Oct 96, Dec 98, Dec 01; Station 7049—Jul 90; Station 7080—Feb 86; Station 7090—Feb 98; Station 7169—Dec 93, Dec 98, Nov 03, Dec 03; Station 7197—Dec 98; Station 8141—Dec 96; Station 8147—Feb 96; Station 9519—Jan 02; Station 10037—Aug 02; Station 10041—Dec 94; Station 10091—Nov 99; Station 10537—Apr 87; Station 10541—Jul 91; Station 10611—Jan 98, Feb 98; Station 10636—Aug 86, Jul 92, Jan 03; Station 10658—Jul 03; Station 10668—Jan 90; Station 10795—Jan 98

[b] Station latitude

[c] Station longitude

[d] Station elevation

[e] Rainfall mean

[f] Rainfall standard deviation

[g] Rainfall maximum

[h] Percentage of zero rainfall months

[i] Correlation dimension

visualization of these observations, Fig. 2 presents a graphical representation of these four statistics of rainfall data from the 62 stations.

There are also significant differences in the 'relationships' among these 62 stations and rainfalls observed therein. For instance, Fig. 3 shows the 'correlation' of rainfall observed at each station with every other (Fig. 3a), the 'proximity' of each station with respect to every other (Fig. 3b), and their combination, i.e. correlation *versus* distance (Fig. 3c). As seen, there are almost perfect correlations between rainfall observed at some stations (not necessarily matching in individual months, but overall matching), but there are also cases where no correlations between rainfall observed at some other stations. Similarly, while some stations are located extremely close (even less than about 50 km), some others are as far as about 2,500 km away. Although there is a general inverse relationship between rainfall correlation and station distance (Fig. 3c), it is not always the case, especially for medium distances (stations neither too close nor too far).

Finally, it is also relevant note that, in addition to spatial variability, significant variability in rainfall is observed for any one of the 62 stations with respect to time (months, seasons, etc.); see, for example, Figs. 4a and 5a for rainfall from two selected stations (2014 and 2019). However, no attempt is made in this study to address this within-station rainfall variability with respect to time period or season or temporal scale.

3 Correlation dimension method

The dimension of a time series is, in a way, a representation of the number of variables dominantly governing the underlying system dynamics. Correlation dimension is a measure of the extent to which the presence of a data point affects the position of the other points lying on the attractor

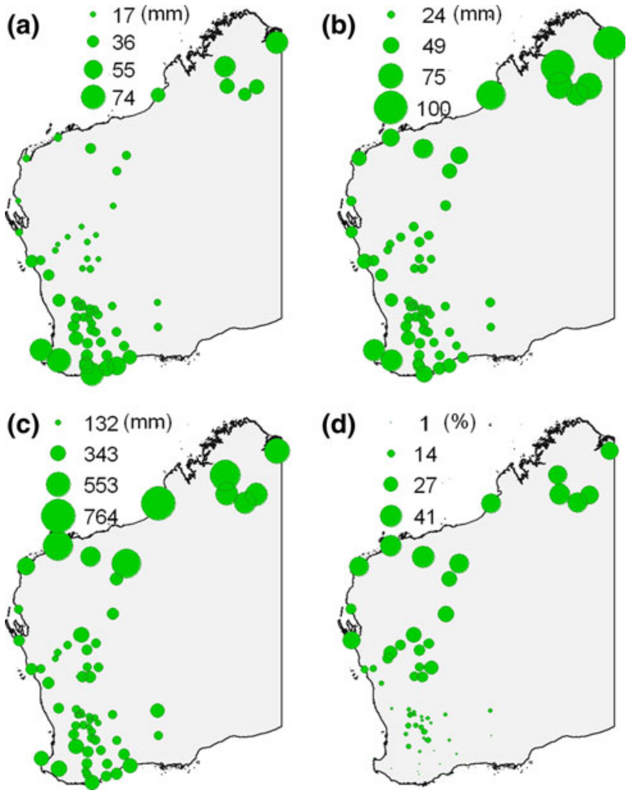


Fig. 2 Statistics of monthly rainfall data at 62 stations in Western Australia: **a** mean, **b** standard deviation, **c** maximum, and **d** percentage number of zeros

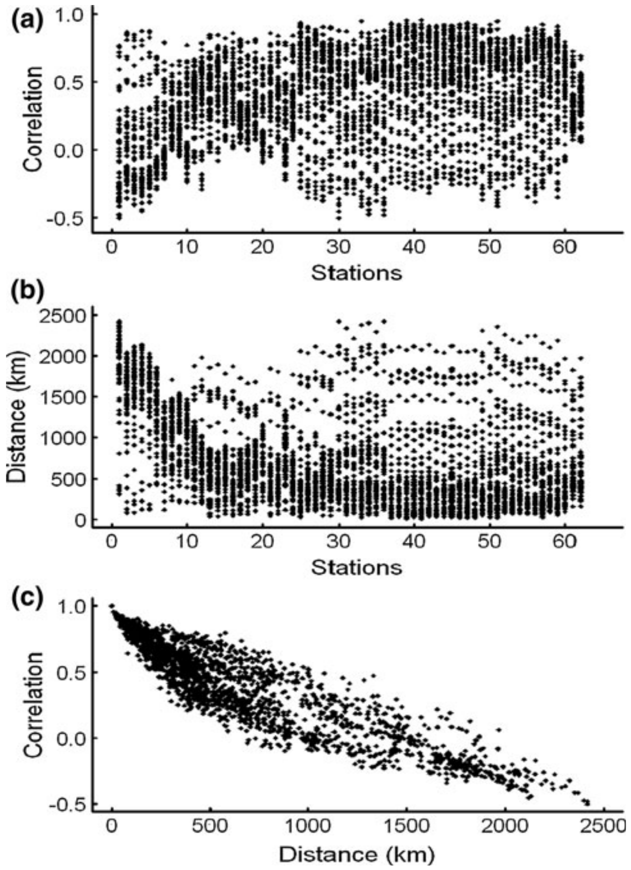


Fig. 3 Relationship among 62 raingage stations in Western Australia: **a** rainfall correlation, **b** station distance, and **c** rainfall correlation *versus* station distance

(where the system eventually settles down to) in the phase space. Phase space is essentially a graph or a co-ordinate diagram, whose co-ordinates represent the variables necessary to describe the state of the system at any moment. The trajectories of the phase space diagram describe the system evolution from some initial state and, hence, represent the history (Packard et al. 1980). The region of attraction of these trajectories provides useful qualitative information on the nature of the system dynamics, such as extent of variability or complexity (see, for example, Sivakumar et al. (2007) for details).

A very common way to reconstruct the phase space is by using the method of delays (Takens 1981). Given a single-variable (or multi-variable) series X_i, where $i = 1, 2, \ldots, N$, a multi-dimensional phase space can be reconstructed as:

$$\mathbf{Y}_j = \left(X_j, X_{j+\tau}, X_{j+2\tau}, \ldots, X_{j+(m-1)\tau}\right) \tag{1}$$

where $j = 1, 2, \ldots, N-(m-1)\tau$; m is the dimension of the vector $\mathbf{Y}_j$, called embedding dimension; and τ is an appropriate delay time (an integer multiple of sampling time). A correct phase space reconstruction in a dimension m generally allows interpretation of the underlying system dynamics in the form of an m-dimensional map f_T, given by:

$$\mathbf{Y}_{j+T} = f_T\left(\mathbf{Y}_j\right) \tag{2}$$

where $\mathbf{Y}_j$ and $\mathbf{Y}_{j+T}$ are vectors of dimension m, describing the state of the system at times j (current state) and $j + T$ (future state), respectively.

The correlation dimension method uses the correlation integral (or function) for determining the dimension of the attractor and, hence, for identifying the extent of complexity of system dynamics (including distinguishing between low-dimensional and high-dimensional systems). Among the many algorithms available for correlation dimension estimation, the Grassberger–Procaccia algorithm (Grassberger and Procaccia 1983) has been the most popular. For an m-dimensional phase space, the correlation function $C(r)$ is given by

$$C(r) = \lim_{N\to\infty} \frac{2}{N(N-1)} \sum_{\substack{i,j \\ (1 \le i < j \le N)}} H\left(r - \left|\mathbf{Y}_i - \mathbf{Y}_j\right|\right) \tag{3}$$

where H is the Heaviside step function, with $H(u) = 1$ for $u > 0$, and $H(u) = 0$ for $u \le 0$, where $u = r - \|\mathbf{Y}_i - \mathbf{Y}_j\|$, r is the vector norm (radius of sphere) centered on $\mathbf{Y}_i$ or $\mathbf{Y}_j$.

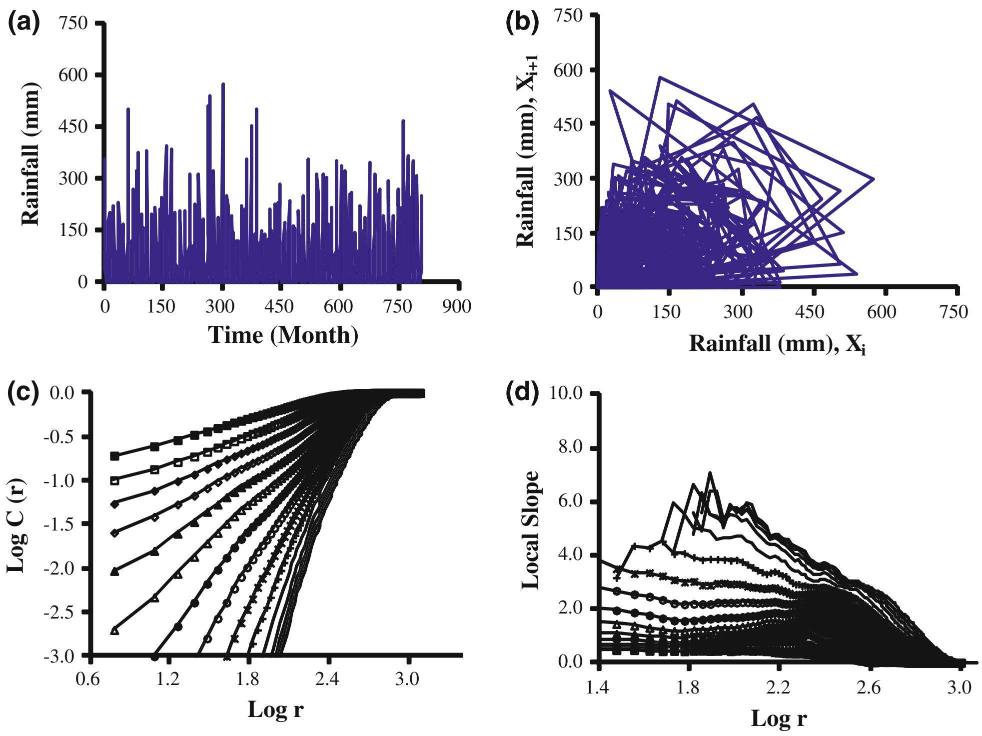


Fig. 4 Analysis of rainfall from Station 2014: **a** time series, **b** phase space, **c** correlation dimension: log C(*r*) versus log *r*, and **d** correlation dimension: local slope

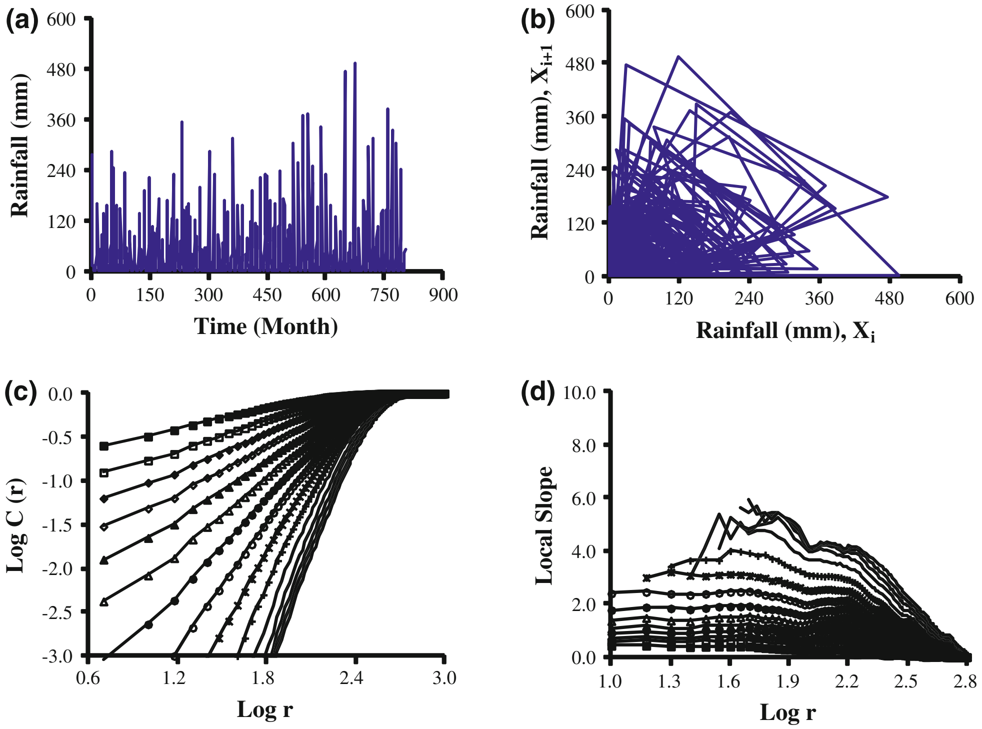


Fig. 5 Analysis of rainfall from Station 2019: **a** time series, **b** phase space, **c** correlation dimension: log C(*r*) versus log *r*, and **d** correlation dimension: local slope

If the time series is characterized by an attractor, then $C(r)$ and r are related according to:

$$\underset{\substack{r\to 0\\ N\to\infty}}{C(r)} \approx \alpha r^{\nu} \tag{4}$$

where α is a constant and ν is the correlation exponent or the slope of the Log $C(r)$ versus Log r plot. The slope is generally estimated by a least square fit of a straight line over a certain range of r (scaling regime) or through estimation of local slopes between r values.

The dimensionality of the time series is determined by checking whether or not there is a saturation of v with increasing m; the saturation value of v is defined as the correlation dimension (d) of the attractor. A low saturation value of v is generally considered to be an indicator of a low-dimensional system, while a high (or no) saturation value is generally considered as an indication of a high-dimensional one. The nearest integer above the saturation value of v, when it occurs, is generally an indication of the number of variables dominantly governing the system dynamics.

There have been many criticisms on the use of the correlation dimension method, especially the Grassberger–Procaccia algorithm, for estimation of dimension of real time series from real systems. These criticisms have been in regards to data size, data noise, presence of zeros, temporal correlations, and others (e.g. Smith 1988; Osborne and Provenzale 1989; Nerenberg and Essex 1990; Theiler 1990; Tsonis et al. 1994; Schreiber and Kantz 1996). As these issues are highly relevant to hydrologic systems, there have been criticisms on the correlation dimension estimates reported for hydrologic data (e.g. Ghilardi and Rosso 1990; Wang and Gan 1998; Schertzer et al. 2002; Koutsoyiannis 2006). These issues as well as interpretations of the dimension estimates in the context of hydrologic data are already available in the literature (e.g. Sivakumar et al. 1999a, b, 2002a, b, 2006; Sivakumar 2000, 2001, 2005a) and, therefore, details are not reported herein. However, as the issue of data size (i.e. number of data points in time series) is particularly relevant for the 62 rainfall series analyzed in this study, we will highlight this issue in the discussion of correlation dimension results in the next section.

4 Results and discussion

4.1 Correlation dimension results

The correlation dimension method is employed to the monthly rainfall data observed at each of the above 62 stations in Western Australia. Figures 4 and 5, for instance, show the results obtained for two of these stations (2014 and 2019), with Figs. 4a and 5a showing the respective rainfall time series. Figures 4b and 5b show the phase space diagrams, reconstructed according to Eq. (1). The diagrams correspond to the reconstruction in two dimensions ($m = 2$) with delay time $\tau = 1$, i.e. the projection of the attractor on the plane $\{X_i, X_{i+1}\}$. For both time series, the projection neither yields a clear attractor in a well-defined region nor is scattered all over the phase space, suggesting that the rainfall dynamics at these stations may exhibit intermediate level of complexity.

The selection of an appropriate τ for phase space reconstruction and correlation dimension estimation has been under considerable debate, with various methods/guidelines exist in the literature, including autocorrelation function (e.g. Holzfuss and Mayer-Kress 1986), mutual information (e.g. Frazer and Swinney 1986), and correlation integral (Leibert and Schuster 1989). It is also not clear how the value of τ obtained from any of these methods is actually relevant to the physical mechanisms that take place in the system, even for artificial ones (see Sivakumar et al. (2007) for the case of Henon map). For instance, use of the autocorrelation function method and selection of the lag time at which the autocorrelation function first crosses the zero line yield $\tau = 3$ for the above two time series (and 3 or 4 for all the 62 rainfall series). Although $\tau = 3$ (or 4) months in these cases seems to suggest some relevance to seasonality, such an outcome may not be (and are not) consistently achieved for all monthly rainfall (and other hydrologic) data or rainfall at another temporal scale (e.g. daily); see, for instance, Sivakumar et al. (2006). Considering these issues and complications, τ is chosen equal to the sampling (i.e. $\tau = 1$ month) for phase space reconstruction of the 62 rainfall time series in this study. The issue of selection of τ has been extensively discussed in the literature (see, for example, Sivakumar et al. (1999a) and Sivakumar (2000)) and, therefore, further details are not reported herein.

Figures 4c and 5c show the correlation functions (i.e. Log $C(r)$ versus Log r) for the above two rainfall series, and Figs. 4d and 5d show the corresponding correlation exponents (i.e. local slopes). The results are those obtained for embedding dimension values from 1 to 15 (from top to bottom lines in Figs. 4c and 5c, and from bottom to top lines in Figs. 4d and 5d). For both time series, the correlation exponent seems to attain saturation after a certain embedding dimension, although the saturation is not as clear as the one that is normally observed for synthetic time series, perhaps due to the presence of noise (see Sivakumar et al. (1999b) for some details). The saturation value (i.e. correlation dimension, d) is approximately 5.44 for data from Station 2014 and 4.63 for data from Station 2019, suggesting that the monthly rainfall dynamics at these stations are dominantly governed by 6 and 5 variables, respectively. Table 1 presents the correlation dimension estimates for all the 62 rainfall time series analyzed. The correlation dimension values are found to be ranging from 4.63 to 8.29. Figure 6 presents a better visualization of the variability of the correlation dimension values of rainfall time series observed across the 62 stations in Western Australia (also roughly grouped according to different ranges of dimension values).

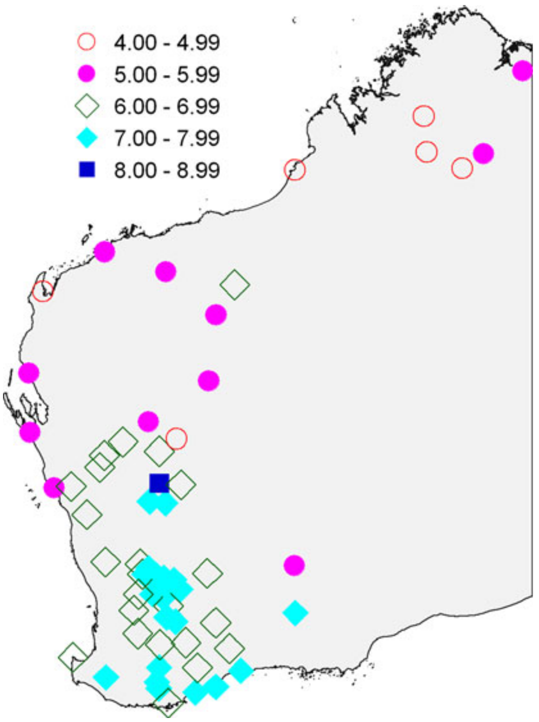


Fig. 6 Correlation dimension results for rainfall across 62 stations in Western Australia

4.2 Rainfall spatial variability

The correlation dimension results presented in Fig. 6 clearly indicate that the rainfall dynamics are far more complex in the southwest part of Western Australia ($d > 6$) compared to the other parts of the state; even within the southwest, the far south exhibits greater complexity in rainfall dynamics ($d > 7$). As for the rest of the state, the mid-western region seems to show slightly more complex rainfall dynamics ($5 < d < 6$) than that in the northeast ($d < 5$). These observations clearly suggest the utility of the correlation dimension method for identification of rainfall variability at point locations and rainfall patterns across space. The fact that the dimension estimate provides useful information on the number of variables dominantly governing the underlying system dynamics is certainly an advantage not only to identify/develop appropriate models but also to find proper ways to couple/integrate the dimensionality (and other nonlinear) concepts with other concepts (e.g. linear).

To be more confident in the above observations and in providing general interpretations, the present results need to be supported further through analysis of rainfall data from far more number of stations and also by employing different methods. As for the number of stations, the situation in the mid-western and northeast parts (and indeed the entire eastern region) of the state is a good case in point. In these parts, the raingage network is very sparse compared to the much denser network that exists in the southwest. The 62 stations considered in the present study have been chosen with due consideration for longer records and no/minimum missing data. Some relaxation to these conditions could yield a significantly larger number of stations for analysis. As for the methods, although the correlation dimension method is a useful and reliable indicator of system dynamic complexity, its potential limitations (as has already been discussed in the literature) calls for additional methods to verify, and possibly confirm, the present results. These methods may include other nonlinear dynamic methods as well as linear methods that can supplement and complement the dimensionality, complexity, and other concepts relevant to rainfall dynamics. We will report the details of such studies in the future.

4.3 Reliability of dimension estimate

A common criticism in the application of the correlation dimension method (especially the Grassberger–Procaccia algorithm) to hydrologic time series is that it significantly underestimates the dimension when the data size is small (e.g. Ghilardi and Rosso 1990; Schertzer et al. 2002). An important basis for this criticism is the assumption of a direct relationship between embedding dimension (m)/correlation dimension (d) and minimum data size (N_{min}); for example, $N_{min} = 42^{\,m}$ or $N_{min} = 10^{2+0.4m}$ or other (e.g. Smith 1988; Nerenberg and Essex 1990). Several studies have responded to the criticisms and concerns regarding the effects of data size on dimension estimate (e.g. Sivakumar et al. 2002a, b, Sivakumar 2005a). These studies emphasize that: (1) there is no direct relationship between m or d and N_{min}; (2) it is more important to assess if the time series is long and representative enough, in terms of period of coverage and sampling time, to capture the essentials of system evolution than to simply look at the data length in terms of the sheer number of values; and (3) if the temporal scale is appropriately chosen, even a few hundred data can provide a reliable dimension estimate.

The correlation dimension estimates for the 62 monthly rainfall time series from Western Australia indeed offer further support regarding the reliability of the correlation dimension method even when the data size is 'small.' With only 804 values in each rainfall series, correlation dimension estimates ranging from as low as 4.63 to as high as 8.29 are obtained, which would be impossible to achieve were the above relationships to hold good. The primary reason for this is that the rainfall data studied herein are long enough (67 years at monthly scale) to adequately represent the dynamic changes that occur in the system evolution. These rainfall series are also significantly longer than the time series (about 300 values) considered sufficient by Sivakumar (2005a), through analysis of stochastic, chaotic, and monthly streamflow series. It is clear, therefore, that the concept of dimensionality (and nonlinear dynamic concepts in general) is indeed useful and effective for studying monthly rainfall variability. Nevertheless, examination of the possible effects of still other factors associated with these rainfall data (e.g. data noise, temporal correlations, presence of zeros) would help gain more confidence in the results obtained and interpretations made. We will attempt this in our future studies. While doing so, we will also delve into the relationship between correlation dimension and coefficient of variation, especially with respect to the above factors.

5 Conclusions

This study investigated the rainfall dynamics across Western Australia using ideas gained from nonlinear dynamic theories. Application of the concept of dimensionality (correlation dimension method) to monthly rainfall data from 62 gaging stations across the state yields interesting results both on rainfall variability at point locations and on rainfall patterns across space. The results indicate that: (1) rainfall dynamics in the state exhibit medium-level complexity; and (2) rainfall dynamics are generally more complex in the southwest when compared to the rest of the state, and the mid-western part shows slightly more complex rainfall dynamics than the northeast.

The present study focused on estimating the level of complexity of rainfall dynamics towards identifying the appropriate level of complexity of model, rather than attempting to make distinction between chaotic and stochastic behaviors towards identifying the type of model. Nevertheless, the dimension estimates offer some useful clues as to the (nonlinear and chaotic) dynamic nature of rainfall as well, although further evidence is needed to support this interpretation. To this end, in our future studies, we will employ both nonlinear and linear methods, to supplement and complement the present analysis.

In the context of nonlinear dynamic and chaos theories for rainfall, the present study is certainly far more extensive than most previous ones in terms of examination of rainfall variability in spatial extent. Although analysis of data from as many as 62 raingage stations can allow reliable interpretations, it is necessary to study a far more extensive network of raingages to be more confident of the interpretations. We will look into this in great detail in our future studies, not only in the context of Western Australia but also for the whole country. The outcomes of such studies could lead to a much better mapping of rainfall variability across Australia. They could also facilitate 'classification' of rainfall regimes and eventually lead to better catchment modeling and water planning and management. Some recent studies could also shed some light to advance research in these directions. For instance, the study by Sivakumar and Singh (2012) presents a simple classification approach (for streamflow) based on correlation dimension estimate, and the study by Sivakumar (2005b) highlights the important role of thresholds (i.e. points at which 'transitions' occur) and their identification.

Acknowledgments This study is supported by the Australian Research Council (ARC). Bellie Sivakumar acknowledges the financial support from ARC through the Future Fellowship grant (FT110100328).

References

Australian Bureau of Statistics (2008) Population projections: Australia. Australian Bureau of Statistics, Canberra

Beesley CA, Frost AJ, Zajaczkowski J (2009) A comparison of the BAWAP and SILO spatially interpolated daily rainfall datasets. In: 18th World IMACS/MODSIM Congress, Cairns, Australia

Berndtsson R, Jinno K, Kawamura A, Olsson J, Xu S (1994) Dynamical systems theory applied to long-term temperature and precipitation time series. Trends Hydrol 1:291–297

Chiew FHS, Young W, Cai W, Teng J (2010) Current drought and future hydroclimate projections in southeast Australia and implications for water resources management. Stoch Environ Res Risk Assess 25(4):601–612

Cleugh H, Smith MS, Battaglia M, Graham P (eds) (2011) Climate change: science and solutions for Australia. CSIRO Publishing, Victoria

CSIRO and Bureau of Meteorology (2007) Climate change in Australia. Technical Report. www.climatechangeinaustralia.gov.au

Frazer AM, Swinney HL (1986) Independent coordinates for strange attractors from mutual information. Phys Rev A 33(2): 1134–1140

Ghilardi P, Rosso R (1990) Comment on "Chaos in rainfall". Water Resour Res 26(8):1837–1839

Grassberger P, Procaccia I (1983) Measuring the strangeness of strange attractors. Physica D 9:189–208

Holzfuss J, Mayer-Kress G (1986) An approach to error-estimation in the application of dimension algorithms. In: Mayer-Kress G (ed) Dimensions and entropies in chaotic systems. Springer, New York, pp 114–122

Jayawardena AW, Gurung AB (2000) Noise reduction and prediction of hydrometeorological time series: dynamical systems approach vs. stochastic approach. J Hydrol 228:242–264

Jayawardena AW, Lai F (1994) Analysis and prediction of chaos in rainfall and stream flow time series. J Hydrol 153:23–52

Jeffrey SJ, Carter JO, Moodie KB, Beswick AR (2001) Using spatial interpolation to construct a comprehensive archive of Australia climate data. Environ Modell Softw 16:309–330

Jones DA, Wang W, Fawcett R (2009) High-quality spatial climate data-sets for Australia. Aust Meteorol Ocean J 58:233–248

Koutsoyiannis D (2006) On the quest for chaotic attractors in hydrological processes. Hydrol Sci J 51(6):1065–1091

Kyoung MS, Kim HS, Sivakumar B, Singh VP, Ahn KS (2011) Dynamic characteristics of monthly rainfall in the Korean Peninsula under climate change. Stoch Environ Res Risk Assess 25:613–625

Lavery B, Joung G, Nicholls N (1997) An extended high-quality historical rainfall dataset for Australia. Aust Meteorol Mag 46:27–38

Leibert W, Schuster HG (1989) Proper choice of the time delay for the analysis of chaotic time series. Phys Lett A 141:386–390

Lorenz EN (1963) Deterministic nonperiodic flow. J Atmos Sci 20:130–141

Mehrotra R, Sharma A (2009) Evaluating spatio-temporal representations in daily rainfall sequences from three stochastic multi-site weather generation approaches. Adv Water Resour 32:948–962

Mehrotra R, Sharma A (2010) Development and application of a multisite rainfall stochastic downscaling framework for climate change impact assessment. Water Resour Res 46: W07526

Nerenberg MAH, Essex C (1990) Correlation dimension and systematic geometric effects. Phys Rev A 42(12):7065–7074

Oke AMC, Frost AJ, Beesley CA (2009) The use of TRMM satellite data as a predictor in the interpolation of daily precipitation over Australia. In: 18th World IMACS/MODSIM Congress, Cairns, Australia

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

Packard NH, Crutchfield JP, Farmer JD, Shaw RS (1980) Geometry from a time series. Phys Rev Lett 45(9):712–716

Puente CE, Obregon N (1996) A deterministic geometric representation of temporal rainfall. Results for a storm in Boston. Water Resour Res 32(9):2825–2839

Rodriguez-Iturbe I, De Power FB, Sharifi MB, Georgakakos KP (1989) Chaos in rainfall. Water Resour Res 25(7):1667–1675

Schertzer D, Tchiguirinskaia I, Lovejoy S, Hubert P, Bendjoudi H (2002) Which chaos in the rainfall-runoff process? A discussion on 'Evidence of chaos in the rainfall-runoff process' by Sivakumar et al. Hydrol Sci J 47(1):139–147

Schreiber T, Kantz H (1996) Observing and predicting chaotic signals: is 2% noise too much? In: Kravtsov YuA, Kadtke JB (eds) Predictability of complex dynamical systems. Springer, Berlin, pp 43–65

Sivakumar B (2000) Chaos theory in hydrology: important issues and interpretations. J Hydrol 227(1–4):1–20

Sivakumar B (2001) Rainfall dynamics at different temporal scales: a chaotic perspective. Hydrol Earth Syst Sci 5(4):645–651

Sivakumar B (2004) Chaos theory in geophysics: past, present and future. Chaos Solitons Fractals 19(2):441–462

Sivakumar B (2005a) Correlation dimension estimation of hydrologic series and data size requirement: myth and reality. Hydrol Sci J 50(4):591–604

Sivakumar B (2005b) Hydrologic modeling and forecasting: role of thresholds. Environ Model Softw 20(5):515–519

Sivakumar B (2009) Nonlinear dynamics and chaos in hydrologic systems: latest developments and a look forward. Stoch Environ Res Risk Assess 23:1027–1036

Sivakumar B, Singh VP (2012) Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework. Hydrol Earth Syst Sci 16:4119–4131

Sivakumar B, Liong SY, Liaw CY, Phoon KK (1999a) Singapore rainfall behavior: chaotic? ASCE J Hydrol Eng 4(1):38–48

Sivakumar B, Phoon KK, Liong SY, Liaw CY (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, Kawamura A (2000) Dynamics of monthly rainfall-runoff process at the Göta basin: a search for chaos. Hydrol Earth Syst Sci 4(3):407–417

Sivakumar B, Sorooshian S, Gupta HV, Gao X (2001) A chaotic approach to rainfall disaggregation. Water Resour Res 37(1):61–72

Sivakumar B, Berndtsson R, Olsson J, Jinno K (2002a) Reply to 'which chaos in the rainfall-runoff process?' by Schertzer et al. Hydrol Sci J 47(1):149–158

Sivakumar B, Persson M, Berndtsson R, Uvo CB (2002b) Is correlation dimension a reliable indicator of low-dimensional chaos in short hydrological time series? Water Resour Res 38(2). doi:10.1029/2001WR00333

Sivakumar B, Wallender WW, Horwath WR, Mitchell JP, Prentice SE, Joyce BA (2006) Nonlinear analysis of rainfall dynamics in California's Sacramento Valley. Hydrol Process 20(8):1723–1736

Sivakumar B, Jayawardena AW, Li WK (2007) Hydrologic complexity and classification: a simple data reconstruction approach. Hydrol Process 21(20):2713–2728

Smith LA (1988) Intrinsic limits on dimension calculations. Phys Lett A 133(6):283–288

Takens F (1981) Detecting strange attractors in turbulence. In: Rand DA, Young LS (eds) Dynamical systems and turbulence, Lecture notes in mathematics 898. Springer, Berlin, pp 366–381

Taschetto AS, England MH (2009) An analysis of late twentieth century trends in Australian rainfall. Int J Climatol 29:791–807

Theiler J (1990) Estimating fractal dimension. J Opt Soc Am A 7:1055–1073

Tsonis AA, Elsner JB (1988) The weather attractor over very short timescales. Nature 333:545–547

Tsonis AA, Triantafyllou GN, Elsner JB, Holdzkom JJ II, Kirwan AD Jr (1994) An investigation on the ability of nonlinear methods to infer dynamics from observables. Bull Am Meteor Soc 75: 1623–1633

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

Western AW, Anderson B, Siriwardena L, Chiew FHS, Seed A, Blöschl G (2011) The within-day behaviour of 6 minute rainfall intensity in Australia. Hydrol Earth Syst Sci 15:2561–2579