**Nonlinear Processes
in Geophysics**

Complexity and predictability of daily precipitation in a semi-arid region: an application to Ceará, Brazil

M. E. S. Silva[1], L. M. V. Carvalho[2], M. A. F. da Silva Dias[2,3], and T. de M. B. S. Xavier[4]

[1]Department of Geography, FFLCH, University of São Paulo, Brazil
[2]Department of Atmospheric Sciences, University of São Paulo, Brazil
[3]Centro de Previsão de Tempo e Estudos Climáticos, Instituto Nacional de Pesquisas Espaciais, Brazil
[4]Department of Hydraulic Engineering and Environment, University of Ceará, Brazil

Abstract. Complexity and predictability of daily precipitation in a tropical semi-arid region (Ceará State, Brazil) is assessed by applying entropy concepts. Precipitation regimes in that region depend on several dynamical forcings, the most important being the displacement and activity of the Intertropical Convergence Zone in the Atlantic Ocean. Topography is another important factor that influences the spatial distribution of rainfall in the region. A hierarchical approach based on sequences of events of different lengths is used to estimate complexity of daily precipitation records. It is shown that precipitation in Ceará exhibit more random than periodic sequences, which indicates a large degree of complexity. Nevertheless, there is indication of potentially inherent rules in the precipitation time-series that could ultimately improve prediction on time-scales between 9–11 days. It is suggested that synoptic-scale disturbances (1–8 days) represent important sources of rules in the precipitation regimes in this region.

1 Introduction

Understanding complexity and ultimately predictability of precipitation regimes on time-scales of a few days has enormous social and economic impacts on tropical semi-arid regions such as the State of Ceará, in the Northeast of Brazil. For instance, the state of Ceará has 7 430 661 inhabitants with 29% of the population living in rural areas. Approximately 29% of the Ceará's territory is occupied by pastures (natural and artificial). Temporary and permanent crops and natural vegetation cover approximately 45% of the territory (source: http://www.ibge.gov.br). Therefore, farming represents a significant segment of Ceará's economy, which suffers profound impact from changes in precipitation.

Correspondence to: M. E. S. Silva
(elisasiq@usp.br)

In the tropics, the largest variations in precipitation and cloudiness occur on seasonal time-scales. A dominant diurnal cycle is also observed and day-to-day weather changes occur, overall, due to changes in position and intensity of migratory tropical and extratropical disturbances and quasi-stationary tropical systems (Asnani, 1993). On interannual time-scales precipitation in Ceará is influenced by El Niño and Southern Oscillation (ENSO) (Ropelewsky, 1987) and the Atlantic Dipole (Moura and Shukla, 1981). These external forcings have increased the seasonal predictability in statistical models (e.g., Repelli and Nobre 2004; Mélice and Servain, 2003). Nevertheless, on daily time-scale, the variability of precipitation is much less predictable and dynamical and statistical models in general show low skill (Bengtsson et al., 2005; Bengtsson and Hodges, 2006). This observation motivated the investigation of predictability on short time-scales in the present study.

In general, the tropical atmosphere is convectively unstable in the lower and middle troposphere. Thermodynamic instability of the atmosphere depends on heat and moisture availability, which in turn, can be enhanced by dynamical forcings on a broad range of time-scales (Garstang and Fitzjarrald, 1999). Variations in dynamical and thermodynamical processes cause non-linear responses in the state of the atmosphere. Non-linearity generates irregularity in dynamical systems (Shukla, 1985; Tsonis, 2001). The irregularity can be completely random or show some order. The existence of order in dynamical systems determines the weather predictability limits in the tropics (Tsonis and Elsner, 1989).

To understand the nature of irregularities in dynamical systems and implications for predictability, we need first to define deterministic, completely random and chaotic systems. Deterministic systems present at least some intrinsic predictability, that is, they are not completely random. A completely random system, on the other hand, can only be described by an infinite number of rules. Deterministic chaos is mathematically defined as an almost random state originated

from simple deterministic systems and in nature many dynamical systems are chaotic. The atmosphere is governed by chaotic processes and, therefore, its evolution is confined to specific areas in the state space restricted to attractors (Tsonis and Elsner, 1989). Strictly speaking, apparent random time-variation of a given atmospheric quantity may contain information about the underlying chaotic processes described by subsystems embedded into the dynamical system. An important question to be answered in this context is how to distinguish deterministic chaos from noise.

Elsner and Tsonis (1992) suggest that predicting time series using nonlinear approaches, such as neural networks, provides a good method for differentiating additive noise from deterministic chaos. This can be demonstrated, for example, by the magnitude of the correlation coefficient between neural network prediction and actual data. The Lorenz system shows a rapid drop in magnitude of the correlation coefficient, as function of prediction time. In contrast, the wave-plus-noise shows constant correlation for all time steps. Nevertheless, the authors show that the neural network clearly outperforms the linear auto-regressive model as prediction time increases. Moreover, linear statistical approaches detect periodicity in time and frequency domains based on the linear dependence between successive points and, therefore, may fail in dealing with nonlinear dynamical system. Thus, the examples above indicate the efficiency of non-linear methods and chaos theory compared to linear methods in providing information about the nature of the signal.

An important question in weather and climate studies is the definition of predictability. By assuming that the atmosphere is a deterministic dynamical system, we can consider a priori the existence of predictability. Predictability is a useful concept to define intrinsic natural characteristic of the atmosphere (Tsonis, 1992) and to evaluate the potential of dynamical models in predicting weather. In this study notions of chaos and theory of information are used for the evaluation of predictability of precipitation in a tropical region on daily time-scale.

The use of information theory was motivated by the similarities between weather/climate and communication systems. The main objective of the communication system is to make faithful replica of the input message at the receiving end of the system despite the noise added into the channel (Leung and North, 1990). The information supplied as input message must be defined quantitatively. This can be done by using entropy estimate, where the detailed calculation involves the structure of the language used in the text, as firstly introduced by Shanon (1948). The parallel with the atmospheric variability is that it can be divided into signal and noise contributions, in which the signal can be 'decoded' and quantified by applying any suitable metric. The estimation of entropy of a system is a formal way to describe its complexity.

Complexity in nature may be seen through the optic of chaos theory. In a broad sense, the theory of complexity is one branch of science that focuses on extracting information (or simpler patterns) from dynamic systems. Grassberger (1986) defines many measures of complexity based on entropy concepts. Those approaches have been largely applied in several fields such as communications (Chomsky, 1966), computer sciences (Chaitin, 1965), biological evolution (Tsonis et al., 1989), climate (Elsner and Tsonis, 1993) and other deterministic chaotic systems (Maddox, 1990; Tsonis and Elsner, 1988).

In summary, information theory provides a useful theoretical framework to understand and quantify weather and climate predictability (Leung and North, 1990; Schneider and Griffies, 1999). Entropy is also a good measure of model skill (Leung and North, 1990; Smith et al., 2001; Richardson, 2000, 2001; Roulston and Smith, 2002). Moreover, information theory and entropy, have been previously used to quantify ensemble spread (Stephenson and Doblas-Reyes, 2000). Roulston and Smith (2002) proposed the use of a skill score for assessing probabilistic forecasts based on the information deficit (or ignorance) given a probabilistic forecast. In that work the authors emphasize the advantages of this method over other scores commonly used, such as the Brier Score.

In the present study, complexity and predictability of rainfall related both to high temporal and spatial variability are assessed by using entropy concepts analogously to Elsner and Tsonis (1993). A hierarchical approach to estimate complexity is applied to daily precipitation records from 36 stations covering the state of Ceará. Our objective is to identify spatial patterns of complexity in the region and ultimately provide information about the potential predictability of precipitation on short-time scale (8–13 days). This paper is organized as follows: an overview on the climatology of the state and dataset are presented in Sect. 2. The methodology is discussed in Sect. 3. Results are presented in Sect. 4 and conclusions in Sect. 5.

2 Regional climatology and data

Ceará State is located in the tropics between the coordinates 3° S–8° S and 42° W–37° W. The rainy season occurs between February and May (Alves, 2000) mainly in response to the southward displacement of the Intertropical Convergence Zone (ITCZ). During these months, the ITCZ moves southward from the Equator in response to atmospheric and oceanic seasonal cycles of the Southern Hemisphere. This displacement is responsible for the largest fraction of the annual precipitation observed in this region (Xavier et al., 2000). The rainy season is also modulated by other atmospheric systems on different time-scales. For instance, the northward incursion of cold fronts eventually reaches the southern part of the Ceará State (Kousky and Ferreira, 1981) and generates atmospheric instability, which contributes to
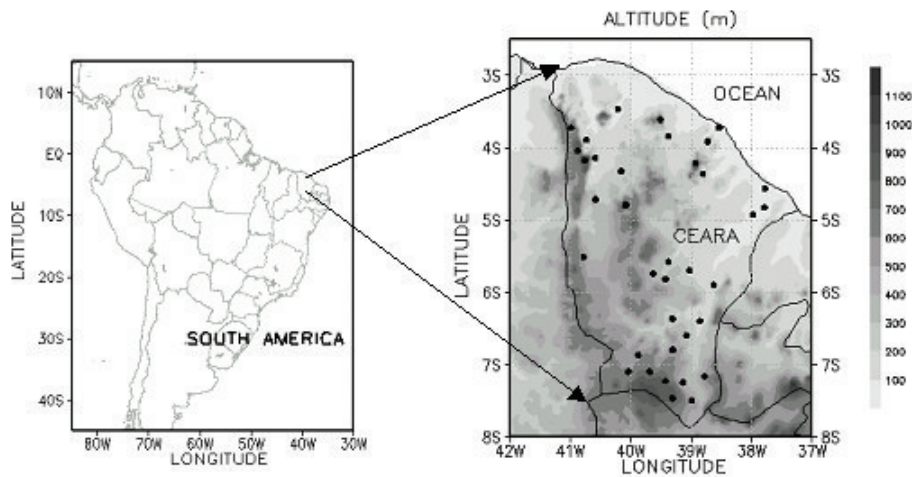
Fig. 1. Topography of Ceará State (shade) and spatial distribution of FUNCEME rain gauge stations (black dots).

increase precipitation in that area. This mechanism seems more important, however, during the pre-rainy season, between November and January. Other variability in subsynoptic time scales, such as the easterly waves (Diedhiou et al., 1998, 1999) seems to contribute to the rainfall regime in this region. On diurnal time-scales, the sea-land breeze plays an important role in modulating precipitation over the costal area (Kousky, 1980).

The combined influence of the systems described above imply on large spatial variability of rainfall in this area. Topography (Fig. 1) is also important to characterize the spatial pattern of precipitation over Ceará. The highest elevations are located at northwest (General Mountain Ridge) and southeast (Araripe Plateau), with average altitudes ranging from 800–900 m and 900–1000 m, respectively. These regions are characterized by high precipitation totals (∼1100 and 800 mm, respectively) during the rainy season (Fig. 2). High topography (800–900 m) is also observed near the coast (Baturité Massive) and is associated with high seasonal (Febraury–May) total precipitation (∼1000 mm). In contrast, the central and southwestern portions of Ceará are relatively dry (semi-arid), with total precipitation below 600 mm (Fig. 2) in the rainy season.

Daily precipitation data were obtained for the period 1974 to 1994 (21 years) from the Meteorological and Hydrological Resources Foundation of Ceará State (FUNCEME) observational network. This data set consists of daily precipitation from 36 meteorological stations distributed over the state, as shown in Fig. 1. Spurious rainfall were checked and eliminated from the data set using the methodology described in Xavier et al. (1995).
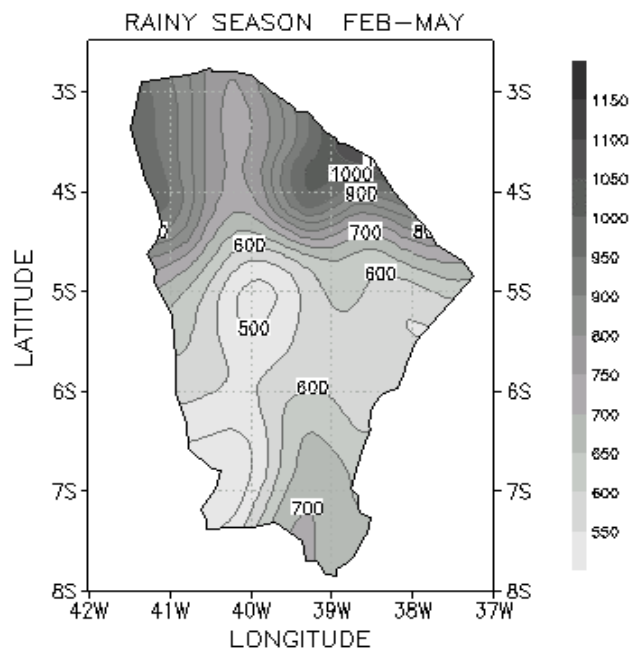

Fig. 2. Climatology of precipitation (mm) (1974–1994) in Ceará during February–May (rainy season).

3 Complexity measures

The challenge of estimating complexity is essentially to find unknown elements (i.e., components, relations, forces, interactions and constraints) existing in dynamical systems (Grassberger, 1986). Intuitively, if there are "n" unknown elements in a system, we can sum them up to get an estimate of complexity. The concept of entropy or information has been part of the classical ideas of complexity (Kolmogorov, 1965; Chaitin, 1965). Entropy is a measure of disorder. Another important concept inherent to the notion of entropy is

that randomness is more complex than periodicity (Elsner and Tsonis, 1993). With such definitions, pure randomness will maximize complexity.

D'Alessandro and Politi (1990) proposed an approach to measure complexity by using a hierarchy of numbers as opposed to a single measure. They demonstrated the usefulness of such definition for describing iterative maps. Elsner and Tsonis (1993) have shown the efficacy of their ideas by measuring complexity of hourly precipitation in 40 years of data. They concluded that, as a result of underlying inherent rules in precipitation formation mechanisms, the observed records are not periodic, but are more complex than a record of random origin. In the current study, the same approach described in Elsner and Tsonis (1993) was adapted to measure complexity in daily precipitation time-series and its spatial variability.

The problem of defining complexity can be reduced by transforming time series in binary sequences of zeros and ones. Consider, for example, two binary sequences (Elsner and Tsonis, 1993),

010101010101010

1001101001011001.

The first sequence can be simply described by repetition of the 01 pattern (or "word"). The second sequence cannot be described by a simple pattern, which is equivalent to say that it is difficult to define a simple rule. The logarithm of the number of words of length n that are found in the sequence (admissible words) as n becomes large is defined as entropy (Maddox, 1990). Words of length n are formed by successively shifting a window of size n through the sequence. Let $N_p(n)$ be the number of possible words of length n ($N_p(n)=2^n$). In the first sequence and for $n=1$ the number of possible words is two, 0 and 1; for $n=2$, the number of possible words is four (00, 01, 11, 10). However, not all combinations occur in the observed time series. Thus, for a given sequence one can define $N_a(n)$ as the number of admissible words of length n contained in the sequence. For instance, in the first sequence and for $n=2$, $N_a(2)=2$ (01, 10).

The first order complexity C^1 can be then defined (D'Alessandro and Politi, 1990) as:

$$C^1 = \lim_{n \to \infty} \log \left[N_a(n) \right]/n. \quad (1)$$

The first order complexity C^1 is a useful notion of complexity and is equivalent to topological entropy. The importance of C^1 is that periodic sequences yield smaller C^1 than non-periodic sequences. However, C^1 fails to recognize random sequence as fairly simple, since C^1 increases as a large number of admissible words are observed in random series. This problem can be solved at the next hierarchical level, after investigating the existence of patterns that do not appear (or are forbidden) in the time-series.

Let's define the words that are forbidden as $N_f(n)$. Thus, for the first sequence, $N_f(2)=2$ (00, 11). Similarly, we can

define forbidden patterns that are irreducible, that is, do not contain any shorter "forbidden words". For instance, for $n=2$ and for the first sequence, the word 00 is forbidden and also irreducible, since it consists of a single admissible word twice (i.e., two "0"). The same is true for the word 11, which consists of two "1". On the other hand, for $n=3$ and for the first sequence, the word "011" is forbidden but not irreducible, whereas the word "000" is forbidden and irreducible.

Therefore, the number of irreducible forbidden words is related to the number of governing rules in the time series: the larger the number of "irreducible forbidden words", the larger the number of rules governing the time series. In other words, computing "irreducible forbidden words" is equivalent to find missing patterns. With these concepts in mind, it is possible to define the second-order complexity as (D'Alessandro and Politi, 1990):

$$C^2 = \lim_{n \to \infty} \log \left[N_{if}(n) \right]/n, \quad (2)$$

where $N_{if}(n)$ is the number of "irreducible forbidden words" of length n. For an infinite and random sequence there should be no forbidden word, so C^2 should be zero, whereas for a chaotic sequence C^2 will in general be greater than zero. Because C^2 is related to the existence of rules governing the behavior of time series, it is also linked to predictability. That is, C^2 magnitudes can be used as a scale to separate randomness (low C^2) from deterministic chaos (high C^2). The second order of complexity C^2 is related to the chaotic properties of the underlying attractor via the relation (Elsner and Tsonis, 1993):

$$C^2 = \lambda_+ D/(1 + D), \quad (3)$$

where D is the dimension of the attractor and λ_+ is the positive Lyapunov exponent (D'Alessandro and Politi, 1990).

In this study, we considered a rainy day when the total daily precipitation was ≥ 2 mm. This threshold corresponds approximately to the 60th percentile of daily rainfall distribution (zero precipitation included) for most stations in Ceará. Therefore, the time series were transformed such that days with rain ≥ 2 mm received 1 and days with rain <2 mm received 0. Complexity of rainfall time-series were addressed in two levels: 1) distinguishing periodic from non-periodic time-series; 2) distinguishing chaotic from completely random time-series.

4 Results

We start the discussion on the complex nature of daily precipitation in Ceará by showing the autocorrelation function (ACF) obtained for each station. The ACF is an efficient linear method to detect existing periodicities and provide a good insight about the random nature of the process (e.g., Chatfield, 1996). In the present study, the autocorrelation was
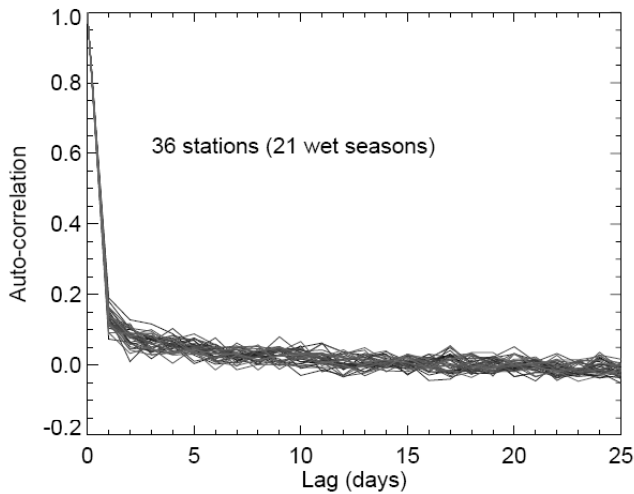
Fig. 3. Average auto-correlation obtained in 21 wet seasons for each station.

Fig. 4. Growth rate of the logarithm of the number of admissible words (C^1) as function of the word length n for daily precipitation records from 36 stations (solid lines) and for a periodic time-series (dashed line).

computed every season and then averaged over 21 years. Figure 3 shows the resulting correlograms obtained for 36 stations, for lags varying from 0–25 days. The fast decrease to zero of all autocorrelations for lags >0 suggests short memory of daily precipitation in the region. This result reinforces the idea that the processes governing daily precipitation in Ceará are likely non-linear and, therefore, gives basis to the application of a more suitable method to detect the chaotic nature of precipitation.

4.1 The first order complexity

We recall that the first order complexity C^1 is related to entropy, which can be described by the number of admissible patterns of any size. Moreover, C^1 is a good measure to distinguish between periodic and non-periodic time series. The number of admissible words $N_a(n)$ was obtained for each station separately. The growth rate of $N_a(n)$ as a function of the word length n for daily precipitation records is shown in Fig. 4. $N_a(n)$ for a simulated periodic sequence was computed and included for comparison with actual data. Figure 4 clearly indicates that all observed time series show higher number of admissible words than the periodic series for the entire interval of word lengths n. This means that observed precipitation records are more complex than simple periodic time series. Moreover, the periodic series shows a higher rate of decrease in the number of admissible words for short word lengths ($n<8$) comparatively to the actual data. It is also interesting to note that for n longer than 8 days the curves describing the variation of C^1 as a function of n are distinct. This suggests the existence of different precipitation regimes in the region. However, as pointed out before, C^1 alone cannot distinguish complete randomness from deterministic systems. This issue will be addressed by examining the second order complexity in the next section.
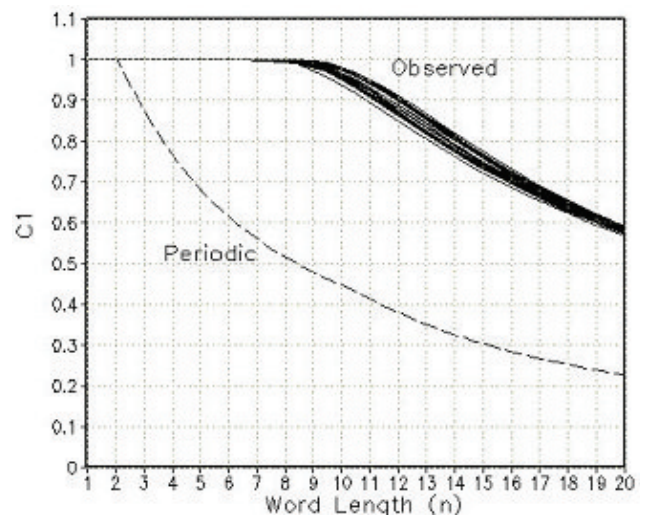
4.2 The second order complexity

As discussed in Sect. 3, the second order complexity C^2 (Eq. 2) distinguishes random from quasi-random (chaotic) systems. We recall that quasi-random systems present more irreducible forbidden words than random systems. Therefore, quasi-random systems are more predictable than random systems, since the former can be described by a higher number of rules. In this study, the second order complexity C^2 was computed for each of the 36 stations. A non-parametric significance statistical test was performed to verify if the computed C^2 for the actual data differs from C^2 obtained from random time series of the same length. In this test, each time series was randomly shuffled 1000 times and C^2 was computed each step. The significance level was assessed based on the frequency distribution of C^2 obtained for the 1001 time series. C^2 was considered statistically significant when it is above the 80th percentile of the resulting distribution. Figure 5 illustrates an example of the frequency distribution of C^2 obtained for N=10 from 1000 randomizations of the precipitation time-series observed at station 27 ("Russas"). The value of C^2 for the observed time-series is indicated in the figure (dashed vertical line) and, in this case, is above the 80th percentile of the distribution, which corresponds to C^2=0,698. The spatial variability of C^2 for all stations is presented in Fig. 6. Areas with statistically significant C^2 (at 80% or higher confidence level) are shaded.

For n=8 days (Fig. 6a), C^2 is relatively small and not statistically significant over the region. As n increases to 9 days (Fig. 6b), C^2 magnitudes increase and statistically significant patterns are observed northward and southward of the state. Notice that the region characterized by semi-arid
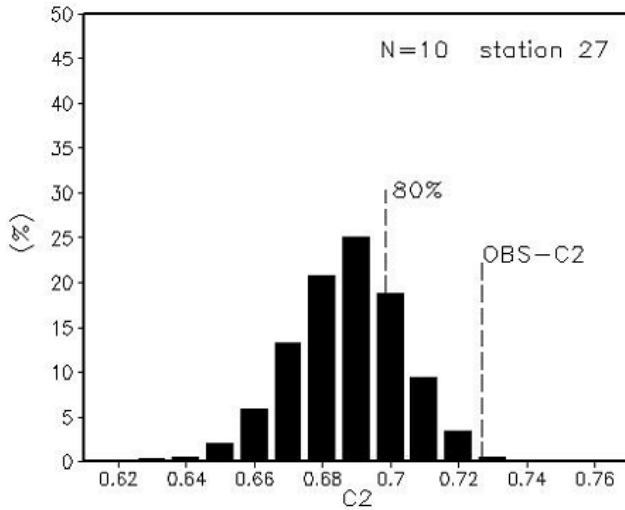
Fig. 5. Frequency distribution of C^2 obtained from 1000 random-izations of the station-27 time-series ("Russas") for N=10. The dashed vertical lines indicate the C^2 80th percentile (=0.698) and the observed C^2 (=0.727).

Fig. 6. Spatial distribution of the number of "irreducible forbidden words" for length N, ranging from 8 to 13 days (**a–f**) in Ceará State, Brazil. Shaded areas indicate statistical significance above 80%.

rainfall regime (central-western Ceará, Fig. 2) shows low values of C^2 for this word length. We recall that the north-ward portion of Ceará is influenced by the variability of the ITCZ convective activity and its interaction with topography (see Fig. 1). In addition, convective activity in the ITCZ ex-hibits a broad range of temporal variability, from synoptic (2–10 days) to interannual (>365 days) time-scales. For in-stance, on synoptic time-scales, easterly waves (4–6 days) from Africa are known to modulate rainfall variability in the northern coast of South America (Diedhiou et al., 1998, 1999). Southern Ceará is also influenced by distinct rain-fall producing systems. Cold fronts (5–10 days time-scale) eventually reach the region and modulate rainfall variability (Kousky, 1980). These disturbances can bring some regular-ity or "rules" to the system and, therefore, according to the concepts of entropy and complexity discussed here, poten-tially increase predictability.

As n increases from 10 to 11 days (Figs. 6c and d), C^2 increases and, except for a very small area, it is statistically significant over most of the state. The behavior of C^2 with respect to the word length suggests that statistically signif-icant C^2 and the word length are related to each other. An example of this relationship for 1000 randomizations of the precipitation observed at station 29 ("Santana do Acaraú") is presented in Fig. 7. Notice that, as n increases, C^2 for the ac-tual data becomes larger than C^2 for the surrogated data. Sta-tistical significance is observed in this particular example for $n>10$. Therefore, long enough words seem necessary to find embedded rules in daily precipitation regimes in Ceará. Nev-ertheless, because we are dealing with limited sample sizes, n cannot increase indefinitely, and one can expect an upper limit in n such that significant results cannot be found. This is
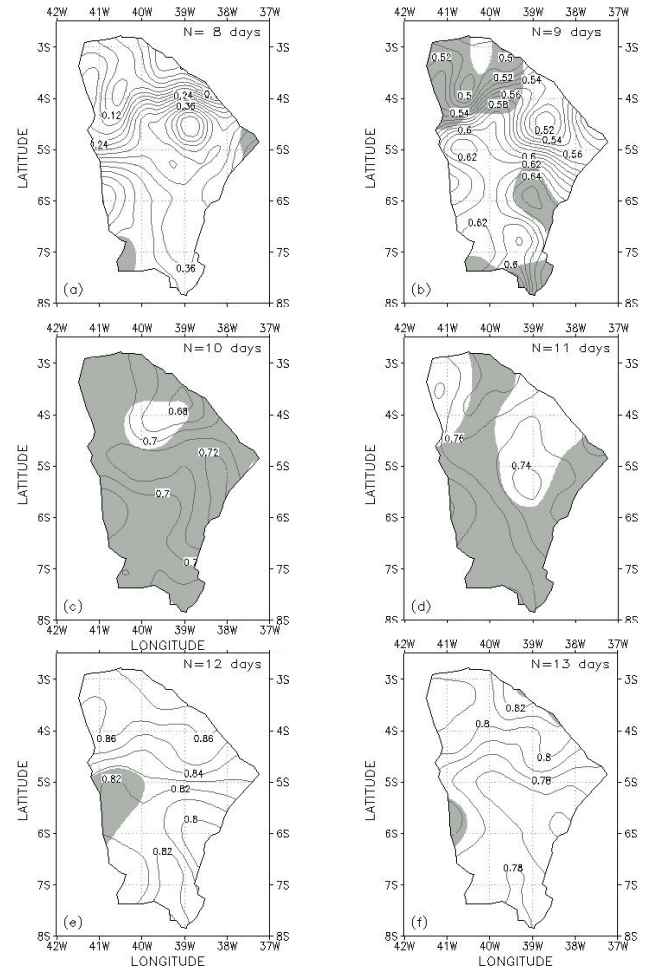
also suggested by Elsner and Tsonis (1993) (see their Fig. 3). In the present case, for $n>12$ days (Figs. 6e and f), statisti-cally significant C^2 values are found in very small areas. We suspect that the lack of statistical significant is likely because the time-series are too short to find embedded rules in large word lengths.

The spatial coherence observed for C^2 might be influenced by spatial correlation among stations. To address this is-sue, the spatial connectivity of the stations is examined (Tso-nis et al., 2006). The basic idea is to determine the maxi-mum correlation obtained between one station and all other 35 stations. With this procedure, the existence of connec-tions or clusters of stations with high spatial correlations are promptly identified. Correlations among stations do not ex-ceed 0.48 (Fig. 8). The average spatial correlation in the re-gion is 0.2, indicating that the stations are poorly connected. Nevertheless, two groups of stations with correlations above 0.4 (Fig. 8) can be identified in the southern and northern
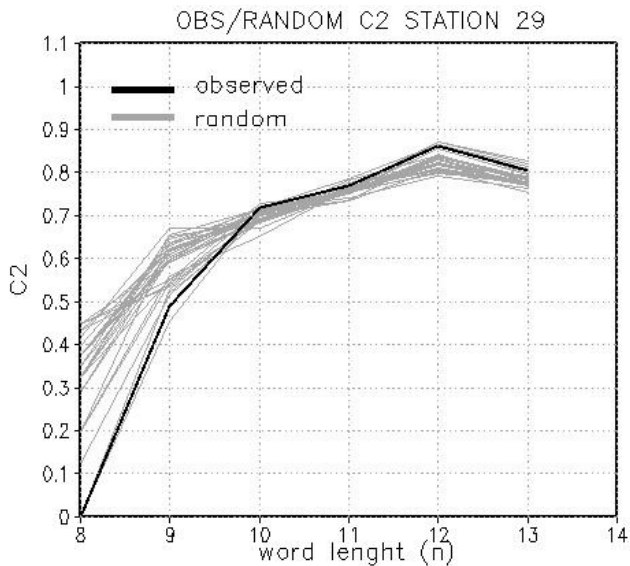
Fig. 7. C^2 for the 1000 randomizations of the station-29 ("Santana do Acaraú") time-series as function of word length. Observed values are indicated by the heavy line. Statistical significance is observed for $n>10$.
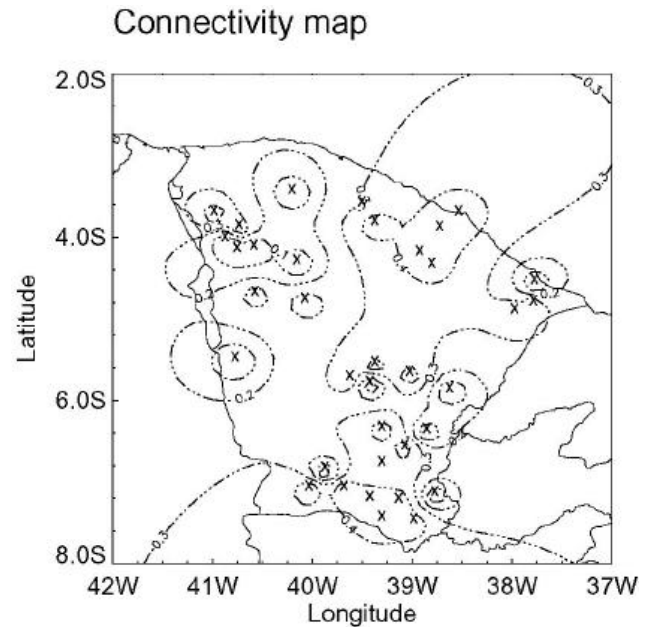

Fig. 8. Connectivity map for 36 stations. Dashed lines show spatial correlation. Station location is indicated by "x".

portions of the region. Topography (Fig. 1) can modulate precipitation in these regions and, therefore, influence connectivity. These results suggest, however, that the statistical significance of C^2 does not seem to be strongly influenced by spatial correlation (compare Fig. 6 with Fig. 8).

To test the hypothesis that synoptic-scale ($\sim$1–10 days) influences the spatial patterns of C^2, a spectral analysis was performed for each station. For this purpose, the spectrum was obtained in each rainy season and then averaged over 21 years. The contribution of the variance of each significant peak of the spectrum (at 95% confidence level) in a given frequency band was then computed. The variance for each frequency band was computed as the area limited between the actual spectrum and the corresponding 95% confidence level curves, assuming either a white or red-noise background spectrum, depending on the lag-1 auto-correlation (Chatfield, 1996). Figure 9 shows the spatial distribution of the variance due to the frequency bands characterized by periods T: $1 \leq T < 4$ days, $4 \leq T < 8$ days; $8 \leq T < 12$ days.

One important aspect of the spatial distribution of the spectral variance shown in Fig. 9 is that the contribution of bands between 1–12 days is statistically significant and not uniform in the region. It can be observed that 1–4 days (Fig. 9a) explains a relatively larger portion of the spectrum relative to the other two bands. For 1–8 days (Figs. 9a–b), the highest contributions are observed in the northern coast and southeastern Ceará, whereas minimum relative contributions are observed in the central-western semi-arid and southern regions. The northern coastal regime is likely modulated by variations in the ITCZ and easterly waves, and interactions

with topography. Conversely, the maximum relative contribution from the 8–12 days band (Fig. 9c) occurs over western and southern Ceará. This suggests the importance of synoptic scale systems (e.g., cold fronts) that propagate from higher latitudes and eventually reach the region, modulating precipitation over southern and western Ceará (Kousky, 1981).

The spectral analysis discussed above showed the importance of synoptic scale variations in the wet season over Ceará. These disturbances add 'rules' to otherwise random systems. They impose non-admissible and irreducible forbidden words and, therefore, reduce randomness for word lengths between 9–11 days. Nevertheless, for shorter world length (N=8 days) there is almost no difference between C^2 from the actual rainfall and random time-series (Fig. 6a). This suggests that the rules imposed by synoptic-scale variations cannot be distinguished for periods shorter than 8 days.

5 Conclusion

Complexity and predictability in a tropical region with large variability of rainfall regimes is addressed in this study. Here, determinism or embedded rules in daily rainfall time series were found by using entropy, complexity and chaos concepts. The region under consideration shows a very complex pattern of seasonal precipitation, which combines multi-scale processes and interaction with topography. It was shown that daily precipitation time-series over Ceará differ from simple periodic signals and present large entropy. By using chaos and complexity notions, we showed that daily precipitation
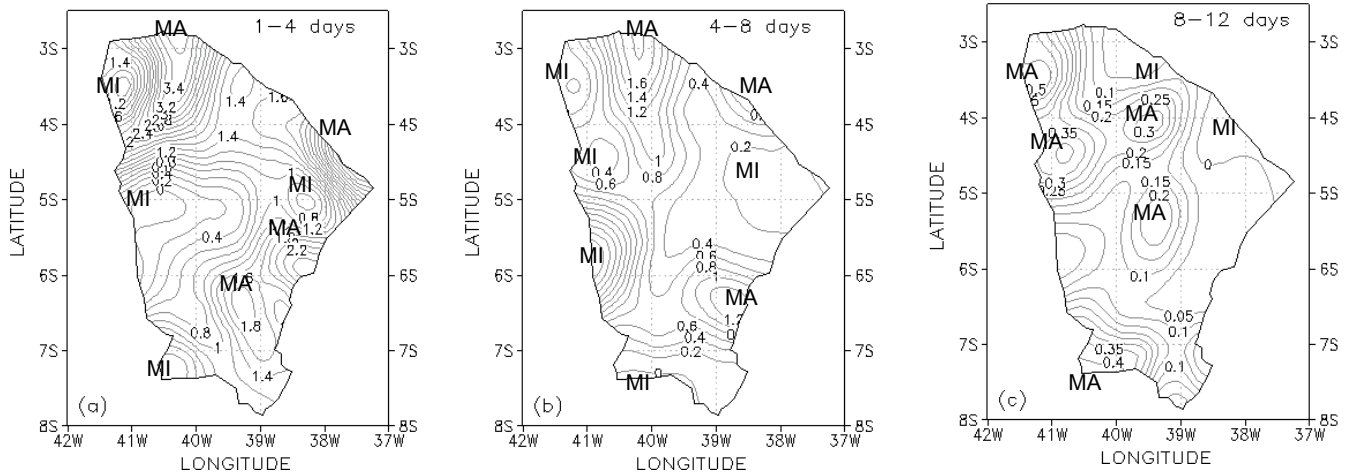
Fig. 9. Spatial distribution of the statistically significant spectral variance (at 95% confidence interval – see text for details) due to bands: 1–4 days **(a)**, 4–8 days **(b)** and 8–12 days **(c)**. Interval between isolines varies according to the band: 0.2 (1–4 and 4–8 days) and 0.05 (8–12 days). Relative maximum (MA) and minimum (MI) variance are indicated.

in the region cannot be considered as the result of completely random processes, suggesting the existence of embedded rules. Statistical linear tools such as autocorrelation are clearly not sufficient to extract information from systems with similar chaotic characteristics. The existence of rules and their relationships with predictability are important for the improvement of weather forecast on short time-scales (9–11 days).

Precipitation predictability also shows spatial variations. For periods shorter than 9 days, complexity and predictability cannot be distinguished from random processes. For N=9 days, the northern, southern and southeastern portions of the state show distinction between random and non-random predictability. For 10–11 days, most of the southern, eastern and western regions, with exception of areas with few stations, show consistent patterns of predictability. Spectral analysis obtained from daily precipitation indicates that the rules observed in 9–11 days might be the result of disturbances with time-scales between 1–8 days. Apparently, the synoptic-scale disturbances do not generate a large number of irreducible forbidden words for periods (word lengths) shorter than 8 days. Longer time-series would be necessary to find rules with word lengths >12 days.

In the context of weather forecast and modeling, this study indicates that to realistically predict rainfall from 9–10 days over Ceará, particularly over eastern and northern regions, further efforts should be done in representing variations on sub-synoptic time-scales (1–8 days) as, for instance, the easterly waves. In addition, it is suggested that interactions with subtropics and, therefore, the ability of predicting the northward propagation of cold fronts (time-scales between 8–12 days) are important aspects for rainfall predictability, particularly in the southern and western semi-arid region. Moreover, interactions of scales and topography are additional factors

that need to be properly represented in models to improve the forecast skill of rainfall on synoptic scales. The relationships with ENSO and the Atlantic Dipole should be also considered. Nevertheless, long periods of available data are necessary to properly investigate this issue using the same methodology. Even so, the present study provides important insights on predictability and can certainly be applied to many purposes, including modeling.

Edited by: J. Kurths
Reviewed by: Y. Gel and A. Tsonis

References

Alves, J. M. B.: Verification of the seasonal prognostics of precipitation in the Ceará State by using percentiles, Rev. Bras. Meteorol., 15, 73–86 (in portuguese), 2000.

Asnani, G. C.: Tropical Meteorology., vol 1., Nobles Printer, India, 604 pp., 1993.

Bengtsson, L., Hodges, K. I., and Froude, L. S. R.: Global observation and forecast skill, Tellus, 57A, 515–527, 2005.

Bengtsson, L. and Hodges, K. I.: A note on atmospheric predictability, Tellus, 58, 154–157, 2006.

Chaitin, G.: On the length of programs for computing finite binary sequences, J. Assoc.Comput. Mach., 13, 547–569, 1965.

Chatfield, C.: The Analysis of Time Series: An introduction, Chapman & Hall, 1996.

Chomsky, N.: Topics in the Theory of Generative Grammar, Mouton & Co., 1966.

D'Alessandro, G. and Politi, A.: Hierarchical approach to complexity with application to dynamical systems, Phys. Rev. Lett., 64, 1609–1612, 1990.

Diedhiou, A., Janicot, S., Vitard, A., and Felice, P.: Evidence of two regimes of easterly waves over West Africa and the tropical Atlantic, Geophys. Res. Lett., 25, 2805–2808, 1998.

Diedhiou, A., Janicot, S., Vitard, A., Felice P., and Laurent, H.: Easterly wave regimes and associated convection over West Africa and tropical Atlantic: Results from the NCEP/NCAR and ECMWF reanalyses, Clim. Dyn., 15, 795–882, 1999.

Elsner, J. B. and Tsonis, A. A.: Nonlinear Prediction, Chaos and Noise, Bull. Am. Meteorol. Soc., 73, 49–60, 1992.

Elsner, J. B. and Tsonis, A. A.: Complexity and Predictability of Hourly Precipitation, J. Atmos. Sci., 50, 400–405, 1993.

Grassberger, P.: Toward a quantitative theory of self-generated complexity, Int. J. Theor. Phys., 25, 907–938, 1986.

Gargstang, M. and Fitzjarrald, D. R.: Observations of surface to atmosphere interactions in the tropics, Oxford University, 1999.

Kolmogorov, A. N.: Three approaches to the definition of the concept "Quantity of information", Prob. Inf. Transm., 1, 1–7, 1965.

Kousky, V. E.: Diurnal Rainfall Variation in Northeast Brazil, Mon. Wea. Rev., 108, 488–498, 1980.

Kousky, V. E. and Ferreira, N. J.: Interdiurnal surface pressure variations in Brazil: their spatial distribution, origins and effects, Mon. Wea. Rev., 109, 1999–2008, 1981.

Leung, L. Y. and North, G. R.: Information theory and climate prediction, J. Climate, 3, 5–14, 1990.

Maddox, J.: Complicated measures of complexity, Nature, 344, 705, 1990.

Mélice, J.-L. and Servain, J.: The tropical Atlantic meridional SST gradient index and its relationships with the SOI, NAO and Southern Ocean, Clim. Dyn., 20(5), 447–464, 2003.

Moura, A. D. and Shukla, J.: On the dynamics of droughts in northeast Brazil: Observation, theory, and numerical experiments with a general circulation model, J. Atmos. Sci., 38, 2653–2675, 1981.

Repelli, C. A. and Nobre, P.: Statistical prediction of sea-surface temperature over the tropical Atlantic, Int. J. Climatol., 24, 45–55, 2004.

Ropelewsky, C. F. and Halpert, M. S.: Global and regional scale precipitations associated with El Niño/Southern Oscillation, Mon. Wea. Rev., 115, 985–996, 1987.

Roulston, M. S. and Smith, L.: Evaluating Probabilistic Forecast using Information Theory, Mon. Wea. Rev., 130, 1653–1660, 2002.

Schneider, T. and Griffies, S. M.: A conceptual framework for predictability studies, J. Climate, 12, 3133–3155, 1999.

Shanon, C. E.: A mathematical theory of communication, Bell. Syst. Technol. J., 27, 379–423, 1948.

Shukla, J.: Predictability, Advances in Geophysics, in: Atmos. and Ocean Modelling. Part B: Weather Dynamics, edited by: Manabe, S., Academic Press: 87–123, 1985.

Smith, L. A., Roulston, M. S., and Hornberger, J.: End to end ensemble forecasting: Towards evaluating the economic value of the ensemble prediction system, ECMWF Tech. Rep., 336, 2001.

Stephenson, D. B. and Doblas-Reyes, F. J.: Statistical methods for interpreting Monte Carlo ensemble forecasts, Tellus, 52A, 300–322, 2000.

Tsonis, A. A.: The impact of nonlinear dynamics in the atmosphere sciences, Int. J. Bifurcation and Chaos, 11, 881–902, 2001.

Tsonis, A. A. and Elsner, J. B.: The weather attractor over very short time scales, Nature, 33, 545–547, 1988.

Tsonis, A. A. and Elsner, J. B.: Chaos, strange attractors and weather, Bull. Amer. Meteorol. Soc., 70, 14–23, 1989.

Tsonis, A. A., Elsner, J. B., and Tsonis, P. A.: On the dynamics of a forced reaction-diffusion model for biological pattern formation, Proc. Nat. Acad. Sci. USA, 80, 4938–4942, 1989.

Tsonis, A. A., Swanson, K. L., and Roebber, P. J.: What do networks have to do with climate?, Bull. Am. Meteorol. Soc., 585–595, 2006.

Xavier, A. F. S., Xavier, T. M. B. S., Silva Dias, M. A. F., and Toledo, S. R. C.: Análise da Qualidade de Séries Diárias da Pluviometria no Estado de São Paulo do Banco de Dados Hidrométricos do DAEE-CTH, in: Desenvolvimento Sustentável dos Recursos Hídricos, edited by: Clarke, R. T., Chaudhry, F., and Baptista, M., ABRH Publications, no. 1 (in portuguese), November 1995.

Xavier, T. M. B. S., Xavier, A. F. S., Silva Dias, M. A. F., and Silva Dias, P. L.: A Zona de Convergência Intertropical – ZCIT e suas relações com a chuva no Ceará (1964–1998), Rev. Bras. Meteorol. (in portuguese), 15(1), 27–43, 2000.