Entropy theory for analyzing water resources in northeastern region of Brazil

Vicente de P. Rodrigues da Silva, Adelgicio F. Belo Filho, Vijay P. Singh, Rafaela S. Rodrigues Almeida, Bernardo B. da Silva, Inajá F. de Sousa & Romildo Morant de Holanda

To cite this article: Vicente de P. Rodrigues da Silva, Adelgicio F. Belo Filho, Vijay P. Singh, Rafaela S. Rodrigues Almeida, Bernardo B. da Silva, Inajá F. de Sousa & Romildo Morant de Holanda (2015): Entropy theory for analyzing water resources in northeastern region of Brazil, Hydrological Sciences Journal, DOI: 10.1080/02626667.2015.1099789

To link to this article: http://dx.doi.org/10.1080/02626667.2015.1099789

Accepted online: 08 Oct 2015.
Abstract Using the Shannon entropy, the space-time variability of rainfall and streamflow was assessed for daily rainfall and streamflow data for a 10-year period from 189 stations in the northeastern region of Brazil. Mean values of marginal entropy were computed for all observation stations and entropy maps were then constructed for delineating annual and seasonal characteristics of rainfall and streamflow. The Mann-Kendall test was used to evaluate the long-term trend in marginal entropy as well as relative entropy for two sample stations. The marginal entropy values of rainfall and streamflow were higher for locations and periods with the highest amounts of rainfall. The entropy values were higher where rainfall was higher. This was because the probability distributions of rainfall and resulting streamflow were more uniform and less skewed. The Shannon entropy produced spatial patterns which led to a better understanding of rainfall and streamflow characteristics.
throughout the northeastern region of Brazil. The total relative entropy indicated that rainfall and streamflow carried the same information content at annual and rainy season time scales.

**Key words** Relative entropy; Mann-Kendall test; Variability; Information transfer

1 INTRODUCTION

Shannon (1948) introduced the concept of entropy as a measure of information, disorder or uncertainty. The Shannon entropy has since been employed in numerous areas (Singh and Rajagopal 1987), such as mathematics (Dragomir et al. 2000), economics (Kaberger and Mansson 2001), ecology (Ricotta 2002), climatology (Kawachi et al. 2001), medicine (Montaño et al. 2001) and hydrology (Singh 1997). An interesting application of entropy has been for reducing the gap between information needs and data collected by monitoring networks (Krstanovic and Singh 1993a,b; Al-Zahrani and Husain 1998; Agrawal et al. 2005; Chen et al. 2007). In this application, stations are evaluated by the transmission of information to and from stations (Markus et al. 2003). Likewise, entropy has been used for assessing the space variability of rainfall, one of the primary constraints to water resources development and water use practices (Silva et al. 2003; Mishra et al. 2009). The main point here is to measure the disorder or uncertainty of the occurrence of rainfall by entropy (Maruyama et al. 2005).

Another entropy concept is one of cross-entropy which has been employed for measuring synchronization between time series (Xie et al. 2010). Altiparmak and Dengiz (2009) used cross-entropy for design of communication networks. In a similar vein, You and Wood (2005) used cross-entropy for assessing the spatial distribution of crop production within geopolitical units for Sub-Saharan Africa. In a recent study, Chen et al. (2007) analyzed the optimum spatial distribution and the minimum number of rain gauge stations for
the Shimen Reservoir in Taiwan. They found that only seven rain gauge stations were needed to provide the necessary information in the region. These studies suggest that the variability of rainfall can be appropriately measured by the Shannon entropy, and hence rainfall characteristics of 1-day resolution time series can be described (Kawachi et al. 2001). The relative entropy, or Kullback-Leibler divergence, is also often used in a variety of contexts as a measure of the discrepancy of two distributions \( p \) and \( q \) (Cover and Thomas 1991). Also, calculating entropy involves estimates of probability density functions, which can be effectively accomplished using kernel density methods (Lake 2009).

When evaluating the rainfall variability in Paraíba state, Brazil, Silva et al. (2003) observed that the entropy of rainfall was higher in localities and periods with high rainfall values, while the minimum values of entropy occurred during the dry season and in low intensity rainfall areas. Mishra et al. (2009) have used the entropy theory to examine the spatial and temporal variability of precipitation time series for the state of Texas, USA. In their study, marginal entropy has been employed to elucidate the variability associated with monthly, seasonal, and annual time series. They showed that the disorder in rainfall amount and the number of days with rain shows a strong spatial gradient and could be related to significant historical drought periods. Brunsell (2010) has utilized information theory metrics to determine the spatial and temporal variability of daily precipitation over the continental United States. He demonstrates the usefulness of an information theory approach to assessing the dynamics of precipitation. For Singh (2011) the entropy theory, comprising the Shannon entropy, seems to have much potential that remains yet to be fully exploited. Thus, most of works have mainly focused on the spatial and temporal variability of rainfall using information theory for temperate zones (Sivakumar 2001; Molini et al. 2006; Hsu et al. 2006) while less attention has been given to methodologies that include streamflow in
tropical climate zones, for improving streamflow forecasts and accounting for cumulative sources of uncertainty.

The objective of this study therefore was to quantify, by employing entropy, the space-time variability of rainfall and streamflow in the northeastern region of Brazil; analyze the change in the distribution between rainfall and streamflow with the aid of relative entropy; measure the distance or discrepancy between a posterior distribution and a prior distribution; and assess long-term trends in marginal entropy and relative entropy of annual and seasonal rainfall and streamflow, using the Mann-Kendall test. The remainder of the paper is organized as follows. Section 2 presents an overview of the study area and data collection, as well as the mathematical description of marginal entropy, relative entropy and the Mann-Kendall test; section 3 describes the results and discusses the analysis; and section 4 presents conclusions of the study.

2 MATERIAL AND METHODS

2.1 Study area

The northeastern region of Brazil, bounded to the north and east by the Atlantic Ocean, covers an area of about 1.5 million square kilometers. Approximately 60% of this region is a semi-arid area. The area is inhabited by more than 30 million people and the economy is mainly based on subsistence rainfed crop production. The northeastern region is influenced by several large-scale precipitation mechanisms. The rainy-season occurs between January and June and the dry-season between July and December. The wet-season occurs between March and May and the normal annual rainfall ranges from 400 to 2000 mm (Silva 2004). The region is dominated by semi-arid climate with heterogeneous vegetation cover and the mean air temperature varies between 15 and 33 °C (Silva et al. 2006).
The predominant vegetation type in the basins is tropical thorn forest (caatinga), and the soil is fairly diversified, formed mainly by lithosoils, regosoils, latosoil and sandy soils (Silva, 2004). The intracratonic basins of northeastern Brazil are part of a Cretaceous rift system developed along pre-existing structural trends in the basement during the opening of the South Atlantic Ocean (Silva et al. 2010a). The basement is composed of highly metamorphosed Precambrian rocks (aligned structurally in a northwest–southeast or east–west direction). The predominant rocks are migmatites, granites, gabbros and amphibolites. The main lithologies in the region are clastic rocks, including breccias and conglomerates, sandstones, siltstones, mudstones and shales (Carvalho, 2000).

2.2 Rainfall and streamflow data

While other methods have been used to assess rainfall variability such as wavelet analysis (Smith et al. 2004) and Hurst exponent (Mishra et al. 2009), we explore trend analysis applied to entropy for assessing both uncertainty and disorder levels of rainfall and streamflow variables throughout time. To investigate the space variability of rainfall and streamflow at both annual and seasonal scales, daily time series of rainfall and streamflow recorded at 189 stations for a minimum period of 10 years (from 1995-2004) in the northeastern region of Brazil were analyzed and annual totals of marginal entropy were obtained. Meteorological data consecutively observed over 10 years have not any negative effect on the results, since such period can be used for description of average yearly meteorology in Brazil. The estimate of annual and seasonal entropy was obtained by averaging the entropy values of each station. The average entropy values computed for observation stations were employed to construct the entropy maps in order to delineate rainfall and streamflow characteristics. A similar procedure was reported by Kawachi et al.
In a study designed for delineation of water resources zones in Japan using at least 8 years of rainfall observations.

In order to illustrate the understanding of the behavior of rainfall and streamflow over time, the temporal trend in the entropy time series was analyzed using data from two weather stations in the period of record from 1968 to 2001. These stations are located in the state of Ceará, namely Icó (latitude: 6º24′04″ S; longitude: 38º51′44″ W; altitude: 153.4 m above sea level) and São Luiz do Curu (latitude: 3º40′12″ S; longitude: 39º14′36″ W; altitude: 38.4 m above sea level). The analyses herein are limited to two stations within the basins ranging in size from 8,528 km² (São Luiz do Curu station) and 12,865 km² (Icó station).

Kayano and Andreoli (2004) observed that the decadal (9–14 years) rainfall variations of the northern part of northeastern Brazil are independently linked to the Pacific Decadal Oscillation (PDO) or to the sea surface temperature decadal variations in the tropical South Atlantic. Likewise, cycles less than 11-year in rainfall time series has been observed in all northeastern Brazil, including the central and southern parts of region (Silva et al., 2010b). Thus, a 14-year moving average was used for eliminating high-frequency cycles in both rainfall and streamflow time series. The filtered time series were then subjected to the trend analyses. The wavelet transform and moving average filter methods are shown to be capable of separating synoptic and seasonal components in time series with minimal errors (Eskridge et al., 1997). The moving average filter method is shown to have the same level of accuracy as the wavelet transform method. However, the moving average can be applied to datasets with missing observations and is much easier to use than the wavelet transform method.

2.3 Marginal entropy

A discrete form of marginal entropy of a single variable x is given by (Shannon, 1948):
\[ H(X) = -\sum_{k=1}^{K} p(x_k) \log (p(x_k)) \]  

(1)

where \( H \) is a measure of the information (more information results in lower entropy and vice versa), \( k \) denotes a discrete data interval and \( p(x_k) \) is the probability mass function of variable \( x \). Variable \( X \) can have only \( K \) outcomes. The values of entropy are given in bits because the base of the logarithm was assumed to be equal to 2 and the probability \( p(x_k) \) is based on the empirical frequency of variable \( x \) which can have only \( K \) outcomes. The marginal entropy can be defined as the average information content of a random variable and can be used as a measure of uncertainty. When the historical monthly time series of a station is considered for the calculation of marginal entropy, it gives randomness associated with the entire length of time series. Marginal entropy is useful and can be used for any type of data set, for example, yearly, monthly, seasonal, rainy days, to evaluate the randomness in the time series (Mishra et al. 2009).

2.4 Relative entropy

For two given probability distributions \( \{p_1, p_2, \ldots, p_k\} \) and \( \{q_1, q_2, \ldots, q_k\} \), the Kullback–Leibler directed divergence or relative entropy is obtained as:

\[
\text{RE} = \sum_{i=1}^{n} p_i \ln \left( \frac{p_i}{q_i} \right) = \sum_{i=1}^{n} p_i \ln p_i - \sum_{i=1}^{n} p_i \ln q_i
\]

(2)

where \( p_i \) is the a posteriori probability of occurrence of the \( i \)th outcome, \( q_i \) is the a priori probability, \( p = \{p_i\}, q = \{q_i\} \). The relative entropy (RE) provides a means of measuring the probabilistic “distance” of the probability distribution \( p \) from \( q \). The RE provides a model formulation in cases where some previous information about the probability distribution \( p \) is already known. That information, called a prior, for the unknown probability distribution \( p \) is used as the initial estimation of \( p \). The principle of minimum relative entropy implies that the
estimate of p can be discriminated from q with a minimum of difference (You et al. 2009). Therefore, RE approach lends itself to a measurement of similarity between two different sources. When p and q are the same distribution, the relative entropy will be zero. In general, the relative entropy measures how different the distributions p and q are.

2.5 Mann-Kendall test

The Mann-Kendall non-parametric test (Mann 1945; Kendall 1975) is applied for assessing trends in rainfall and streamflow time series. This test is based on statistic $S$ defined as:

$$S = \sum_{i=2}^{n} \sum_{j=1}^{i-1} \text{sgn}(x_i - x_j),$$

(3)

where $x_j$ are the sequential data values, $n$ is the length of the time series and $\text{sgn}(x_i - x_j)$ is -1 for $(x_i - x_j) < 0$, 0 for $(x_i - x_j) = 0$, and 1 for $(x_i - x_j) > 0$. The mean $E[S]$ and variance $\text{Var}[S]$ of statistic $S$ may be given as:

$$E[S] = 0$$

(4)

$$\text{Var}[S] = \frac{n(n-1)(2n+5)}{18} - \sum_{t} t_r (t_r - 1)(2t_r + 5)$$

(5)

where $t_r$ is the number of ties for the $r$th value and $s$ is the number of tied values. The second term represents an adjustment for tied or censored data. The standardized test statistic ($Z_{MK}$) is computed as:

$$Z_{MK} = \begin{cases} 
\frac{S - 1}{\sqrt{\text{Var}(S)}} & \text{if } S > 0 \\
0 & \text{if } S = 0 \\
\frac{S + 1}{\sqrt{\text{Var}(S)}} & \text{if } S < 0 
\end{cases}$$

(6)
The main reason to use a nonparametric statistical test is that no assumption is necessary on the distributional form of the tested data set, which is important when studying hydrometeorological time series (Ehsanzadeh et al. 2010). The presence of a statistically significant trend was evaluated for testing the null hypothesis that no trend existed. A positive $Z_{MK}$ value indicates an increasing trend while a negative one indicates a decreasing trend. To test for either increasing or decreasing monotonic trend at the $p$ significance level, the null hypothesis was rejected if the absolute value of $Z_{MK}$ was greater than $Z_{1-p/2}$, which was obtained from the standard normal cumulative distribution table. In this study, the significance levels of $p = 0.01$ and $0.05$ were applied. The non-parametric estimate of the magnitude of the slope of trend was obtained as follows (Hirsch et al. 1982):

$$\beta = \text{Median} \left[ \frac{(x_j - x_i)}{(j - i)} \right] \text{ for all } i < j$$

(7)

where $x_j$ and $x_i$ are the data points measured at times $j$ and $i$, respectively. Given the uncertainty in precipitation and streamflow forecasts, the direction of some changes is not known. However, it seems likely that the trend analysis tool in entropy enables to identify the presence or absence of trends in the input dataset. In our study, different rainfall and streamflow time series were considered individually, because it is useful to understand the uncertainty or variability within each time series and compare in terms of their variability (Mishra et al. 2009).

3 RESULTS AND DISCUSSION

3.1 Marginal entropy
Mean values of marginal entropy of rainfall and streamflow at two sample stations in northeastern region of Brazil for the year and dry and rainy seasons are shown in Table 1. For both stations, marginal entropy values of rainfall and streamflow were low during the dry season and high during the rainy season. The values of mean annual entropy were very similar to those for the rainy season.

For Icó station, 89 and 86% of the annual entropies of streamflow and rainfall, respectively, were observed in the rainy season. Similarly, for the São Luiz do Curu station, 71 and 87% of the annual entropies of streamflow and rainfall, respectively, were observed in the rainy season. In general, the coefficient of variation was high, particularly for streamflow at both stations, reaching a maximum value of 267.1% during the dry season at the São Luiz do Curu station and a minimum value of 53.9% at the Icó station during the year. The CV values of the marginal entropy reached a maximum of 114.7% at the São Luiz do Curu station for rainy season rainfall and a minimum of 34.2% at the Icó station for annual rainfall. These results suggest that rainfall is less variable than streamflow in the semi-arid area of the northeastern region of Brazil. For less variable rainfall entropy should be less. For example, if the variance is zero, then entropy will be zero as well. The difference between marginal entropy values of rainfall and streamflow among periods can be interpreted as a characteristic of these variables in the studied region. There is an exponential relationship between coefficient of variation and marginal entropy with high coefficient of determination. This relationship is evident because marginal entropy is also a measure of variability of the time series. The two main measures of uncertainty are the entropy and the variance. However, entropy is more general and has some advantages facing to the coefficient of variation (Lawrence, 1999). The entropy is a concave function allows its use as an uncertainty function and the variance is a measure of dispersion and its simplicity remains its major attraction.
For both stations, mean values of relative entropy were higher for annual and rainy season rainfall than those for dry season rainfall (Table 2). Also, relative entropy values tended toward zero in the dry season, indicating a high degree of synchrony or similarity between rainfall and streamflow during this season while for the year and rainy season they were almost equal. Therefore, the total relative entropy showed that annual and rainy season rainfall and streamflow carried the same information content. On the other hand, CV was higher for the dry season and smaller for the year for both stations. The highest values of marginal entropy of rainfall and streamflow always occurred during drought years. Similar results were obtained by Mishra et al. (2009) who observed in a study of variability in rainfall using entropy that high disorderliness in the amount of rainfall and number of rainy days caused severe droughts during the 1950’s in the entire state of Texas, USA. As relative entropy is a measure of the discrepancy between two probability distributions, their mean values for the year and rainy season indicated that rainfall was qualitatively similar to streamflow. The minimization of the relative entropy provides the definition of optimal updating rules for the reference parameter of the density functions and the generation of improved feasible vectors (Altıparmak and Dengiz 2009).

3.2 Trend in marginal entropy

Temporal course of annual and seasonal values of marginal entropy of streamflow and rainfall for Icó and São Luiz do Curu stations are shown in Figs. 1 and 2, respectively. Table 3 presents slopes and the significance level (p-value) for streamflow and rainfall data, as well as for marginal entropy and relative entropy during the year and rainy and dry seasons at Icó and São Luiz do Curu stations. Despite decreasing trends in annual and rainy season rainfall and streamflow at both stations (Table 3), an increasing trend of marginal entropy in
streamflow and rainfall was observed during the year and rainy and dry seasons (Figure 1). On the other hand, streamflow and rainfall data presented an increasing trend during the dry season at Icó station, but only streamflow time series was statistically significant based on the Mann-Kendall test. This finding suggested that there was an increasing uncertainty in annual streamflow at São Luiz do Curu station and in rainy and dry seasons at Icó station. Also, a reduction in uncertainty in streamflow in the rainy season was observed at São Luiz do Curu station. These results suggested that the temporal trend of entropy was not influenced by the original data. Kawachi et al. (2001) showed that average annual entropy and average annual rainfall were less mutually related with a coefficient of correlation of 0.19. The increasing trend in marginal entropy of annual streamflow at Icó station was statistically significant at the $p<0.05$ significance level. On the other hand, the trend in marginal entropy was statistically significant for annual rainfall at São Luiz do Curu station based on the Mann-Kendall test ($p<0.05$) and for dry season ($p<0.01$). When evaluate the long-term persistence and trend in the variability of precipitation using the Hurst exponent and the Mann–Kendall, Mishra et al. (2009) found a distinct spatial patterns in annual series and different seasons.

3.3. Trend in relative entropy

The relative entropy time series did not show any statistically significant trend at the $p<0.05$ level according to the Mann-Kendall test (Table 3). Fig. 3 depicts annual and seasonal values of relative entropy for Icó and in São Luiz do Curu stations. It is seen that for both stations relative entropy values during the dry season were close to zero or even zero in a few years. These values ranged from zero to 0.40 bits for São Luiz do Curu station and from zero to 0.27 bits for Icó station. The range of relative entropy for the year was similar to that during the rainy season. This means that the rainy season entropy mimicked the yearly entropy range.
and the entropy ranges for other seasons were within these ranges. For instance, relative entropy at Icó station ranged from 0.24 to 1.67 bits for the year and from 0.22 to 1.65 bits for the rainy season. Similarly, these values at São Luiz do Curu station ranged from 0.22 to 1.51 bits for the year and from 0.17 to 1.41 bits for the rainy season. Analyzing rainfall data of both stations, higher values of relative entropy occurred during wet years while smaller values occurred during drought years. Therefore, unlike marginal entropy, RE was not related to high disorderliness in the amount of rainfall as observed in the present study and by Mishra et al. (2009).

3.4 Spatial distribution of marginal entropy

Spatial distributions of entropy in annual and rainy and dry rainfall and streamflow in the northeastern region of Brazil are shown in Figs. 4, 5 and 6, respectively. The isoinformation lines of marginal entropy of annual and dry season rainfall and streamflow were higher throughout the coast east of the region (Figs. 4 and 5). However, higher values of entropy during the rainy season were located in the northern part of the region.

As a natural consequence, higher rain might occur alternately during other periods of the year over the northeastern region. Minimum and maximum values of entropy in rainfall and streamflow were observed in the same area for all analyzed periods. For instance, both marginal entropies of annual rainfall and streamflow were minimum in the central area of northeastern region of Brazil, which corresponded to most of the semi-arid region (Fig. 4). The entropy values of rainfall and streamflow were maximum in eastern and northern areas of the region which corresponded to the most northeastern rainy areas. During the rainy season, the entropy decreased from 5.5 bits in the north to 1.5 bits in the south for rainfall and from 6.0 bits to 2.5 bits for streamflow (Fig. 5). On the other hand, during the dry season
entropy values of rainfall and streamflow reached minimum values as compared to the other two periods as a consequence of the rainfall reduction (Fig. 6). Mishra et al. (2009) also used marginal entropy to investigate the space variability of rainfall time series for the State of Texas, USA. They observed distinct spatial patterns in annual series and different seasons and that the variability of rainfall amount as well as number of rainy days within a year increased from east to west of Texas.

The spatial distribution of marginal entropy was practically uniform during the dry season over almost the entire region, particularly for rainfall, with a mean value about of 0.5-1.5 bits. Martín and Rey (2000) analyzed the role of entropy to provide some mathematical arguments for justifying the use and interpretation of entropy as a measure of diversity and homogeneity. As shown in Fig. 4, the isoinformation lines of rainfall divided the whole study region into two clusters, at left with higher values in entropy and at right with lower values of entropy. Marginal entropy values of streamflow were lower in most semi-arid areas, reaching values of until 0.1 bits during the dry season. Therefore, the marginal entropy of rainfall and streamflow was high in areas and periods with the highest amounts of rainfall. Despite large variations of marginal entropy for rainfall and streamflow between periods and even between stations, the overall analysis showed much less variation of entropy during the dry season due to the lower rain rates. By changing the time scale of observations and repeating the entropy calculations, the variability can be changed from one scale to another. The entropy values were higher where rainfall was higher. This was because the probability distributions of rainfall and resulting streamflow were more uniform and less skewed. The less variation of entropy meant that the probabilistic characteristics were not significantly different between areas and periods.

The advantage of the entropy method is that it provides a measure of information or uncertainty at each station. At each station one can also specify the probability distribution
which may or may not be the same for one station to another. By plotting or constructing
isoinformation lines or contours, the spatial variability is exhibited. This can also shed light
on the adequacy of gauging station network. By changing the time scale of observations and
repeating the entropy calculations, the spatial variability can be quantified from one scale to
another. This paper is different from other studies (Maruyama et al., 2005; Mishra et al.
2009) did.

4 CONCLUSIONS

Entropy leads to a better understanding of time and space structure of rainfall and streamflow
in the northeastern region of Brazil. It is shown that entropy can be effectively used for
assessing rainfall and streamflow variability in both space and time. The uncertainty level in
streamflow data is higher than in rainfall data due to the disorder and randomness of
streamflow records. The uncertainty in streamflow is higher than in rainfall because the
propagation of uncertainty from rainfall to streamflow is also compounded by geomorphic
and land use characteristics. The time-space variability of marginal entropy in streamflow
depends on various factors, including the intensity-duration-frequency of rainfall and their
variability in space and time. Both rainfall and streamflow variability can satisfactorily be
obtained in terms of marginal entropy as a comprehensive measure of the regional
uncertainty of these hydrological events. The Mann-Kendall test suggests that the temporal
trend of entropy in rainfall and streamflow is not influenced by the eventual trend of the
rainfall and streamflow values. Relative entropy analysis shows that rainfall and streamflow
carry the same information content for the year and the rainy season. The relative entropy
varies from values close to zero in the dry season to values less than one bit in the year and
rainy season, indicating a high degree of similarity between rainfall and streamflow,
particularly during the dry season due to low flow originating from rainfall. Based on
observed rainfall and streamflow data, it was developed a diagnostic (rather than prognostic) of temporal trends that can be used to formulate inferences on the climate variability. Further, considering rainfall and streamflow simultaneously, it is possible to determine if rainfall and streamflow carry the same information content at certain time scales.

**Funding** The research was supported from grants provided by National Council for Scientific and Technological Development (CNPq), Brazil, (Grant No. 472535/2006-4).

**REFERENCES**


Table 1 Mean values of marginal entropy and the coefficient of variation (CV) for the year and annual and rainy and dry season rainfall and streamflow for Icó and São Luiz do Curu stations

<table>
<thead>
<tr>
<th></th>
<th>Streamflow</th>
<th>Rainfall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Annual</td>
<td>Rainy</td>
</tr>
<tr>
<td>Marginal entropy (bits)</td>
<td>6.4</td>
<td>5.7</td>
</tr>
<tr>
<td>CV (%)</td>
<td>53.9</td>
<td>108.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Streamflow</th>
<th>Rainfall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Annual</td>
<td>Rainy</td>
</tr>
<tr>
<td>Marginal entropy (bits)</td>
<td>7.5</td>
<td>5.3</td>
</tr>
<tr>
<td>CV (%)</td>
<td>109.5</td>
<td>117.5</td>
</tr>
</tbody>
</table>

Table 2 Mean values of relative entropy (bits) and coefficient of variation (CV) for the year and rainy and dry seasons at Icó and São Luiz do Curu stations

<table>
<thead>
<tr>
<th></th>
<th>Icó station</th>
<th>São Luiz do Curu station</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Annual</td>
<td>Rainy season</td>
</tr>
<tr>
<td>Relative entropy</td>
<td>0.79</td>
<td>0.75</td>
</tr>
<tr>
<td>CV (%)</td>
<td>45.2</td>
<td>46.6</td>
</tr>
</tbody>
</table>
Table 3 Slope and significance levels ($p$-values) of the times series of streamflow ($S$), rainfall ($R$), marginal entropy in $S$, marginal entropy in $R$ and relative entropy during annual period and rainy and dry seasons at Icó and São Luiz do Curu stations

<table>
<thead>
<tr>
<th>Variables</th>
<th>Annual</th>
<th>Rainy season</th>
<th>Dry season</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>$p$-value</td>
<td>Slope $p$-value</td>
</tr>
<tr>
<td>Icó station</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streamflow ($S$)</td>
<td>-13.51</td>
<td>0.900</td>
<td>-29.4 0.723</td>
</tr>
<tr>
<td>Rainfall ($R$)</td>
<td>-0.84</td>
<td>0.952</td>
<td>-0.502 0.741</td>
</tr>
<tr>
<td>Marginal entropy in $S$</td>
<td>0.022</td>
<td>0.030</td>
<td>0.0123 0.219</td>
</tr>
<tr>
<td>Marginal entropy in $R$</td>
<td>0.004</td>
<td>0.787</td>
<td>0.0013 0.667</td>
</tr>
<tr>
<td>Relative entropy (RE -R)</td>
<td>-0.031</td>
<td>0.682</td>
<td>-0.0011 0.631</td>
</tr>
<tr>
<td>São Luiz do Curu station</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streamflow ($S$)</td>
<td>-188.5</td>
<td>0.138</td>
<td>-158.4 0.298</td>
</tr>
<tr>
<td>Rainfall ($R$)</td>
<td>-6.60</td>
<td>0.496</td>
<td>-5.67 0.447</td>
</tr>
<tr>
<td>Marginal entropy in $S$</td>
<td>0.11</td>
<td>0.406</td>
<td>-0.06 0.070</td>
</tr>
<tr>
<td>Marginal entropy in $R$</td>
<td>-0.04</td>
<td>0.003</td>
<td>-0.011 0.733</td>
</tr>
<tr>
<td>Relative entropy (RE -R)</td>
<td>0.00065</td>
<td>0.681</td>
<td>0.0004 0.833</td>
</tr>
</tbody>
</table>
Figure caption list

**Fig. 1** Annual and seasonal values of marginal entropy in streamflow (a) and rainfall (b) at Icó station
Fig. 2 Annual and seasonal values of streamflow (a) and rainfall (b) entropy at São Luiz do Curu station.
Fig. 3 Annual and seasonal values of relative entropy at Icô station (a) and at São Luiz do Curu station (b).
Fig. 4 Spatial distribution of entropy of annual rainfall (a) and streamflow (b) in the northeastern region of Brazil
Fig. 5 Spatial distribution of entropy of rainy season rainfall (a) and streamflow (b) in the northeastern region of Brazil.
Fig. 6 Spatial distribution of entropy of dry season rainfall (a) and streamflow (b) in the northeastern region of Brazil