Science of the Total Environment 544 (2016) 330–338

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Shannon information entropy for assessing space–time variability of rainfall and streamflow in semiarid region

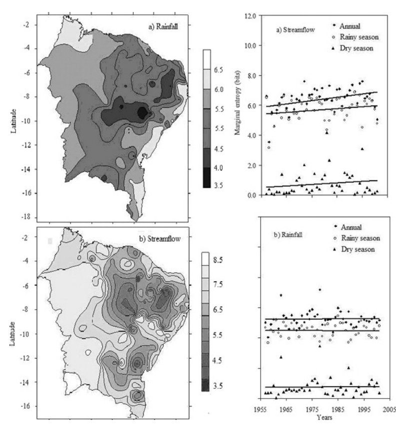
Vicente de P. Rodrigues da Silva [a], Adelgcio F. Belo Filho [a], Rafaela S. Rodrigues Almeida [a], Romildo Morant de Holanda [b], João Hugo Baracuy da Cunha Campos [c]

[a] Federal University of Campina Grande, Av. Aprígio Veloso, 882, Bodocongó, 58109 970 Campina Grande, Paraíba, Brazil
[b] Federal University Rural of Pernambuco, UFRPE, Recife, Pernambuco, Brazil
[c] Estadual University of Paraíba, Campina Grande, Paraíba, Brazil

HIGHLIGHTS

- Entropy can be used for assessing rainfall and streamflow variability
- The uncertainty level in streamflow data is higher than in rainfall data
- Rainfall and streamflow variability can be obtained in terms of marginal entropy
- Rainfall and streamflow carry the same information content

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 29 July 2015
Received in revised form 29 October 2015
Accepted 17 November 2015
Available online xxxx

Editor: D. Barcelo

Keywords:
Mann–Kendall test
Variability
Information transfer
Measure the disorder

ABSTRACT

The principle of maximum entropy can provide consistent basis to analyze water resources and geophysical processes in general. In this paper, we propose to assess the space-time variability of rainfall and streamflow in northeastern region of Brazil using the Shannon entropy. Mean values of marginal and relative entropies were computed for a 10-year period from 189 stations in the study area 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. High degree of similarity was found between rainfall and streamflow, particularly during dry season. 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 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.

E-mail address: vicente@dca.ufcg.edu.br (V.P. Rodrigues da Silva).

http://dx.doi.org/10.1016/j.scitotenv.2015.11.082

1. Introduction

Shannon information entropy introduced by Shannon (1948) has been widely used in numerous areas (Murat, 2012; Asl et al., 2012; Beenamol et al., 2012; Aquino et al., 2013; Bafroui and Ohadi, 2014; Lin and HoInstitute, 2015; Perugini et al., 2015; Aguiar and Guedes, 2015) as a measure of information, disorder or uncertainty. Specifically, Shannon entropy quantifies the expected value of the information contained in a message (Wu et al., 2012). It is considered by many studies as suitable for analyzing of water resources as subject of considerable scientific research in the last decades (Silva et al., 2003; Maruyama et al., 2005; Chen et al., 2007; Mishra et al., 2009; Brunsell, 2010; Liu et al., 2013; Salas and Poveda, 2015).

Entropy theory, comprising the Shannon entropy, seems to have much potential that remains yet to be fully exploited (Singh, 2011). 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.

An interesting application of entropy has been for reducing the gap between information needs and data collected by monitoring networks (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 (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). The relative entropy, or Kullback–Leibler divergence, is also often used in a variety of contexts as a measure of discrepancy of two distributions. Another entropy concept is one of cross-entropy which has been employed for measuring synchronization between time series (Xie et al., 2010).

Most of works, comprising the Shannon entropy, 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 rainfall and streamflow in tropical climate zones for these variables estimates improvement at a time scale, shifting from years to days by exploiting the time series structure. The cumulative sources of uncertainty in rainfall and streamflow remain practically unknown and have not yet been investigated in systematic manner. To address this issue, we used the Shannon entropy to quantify the variability of rainfall and streamflow in the northeastern region of Brazil. Once the time series are obtained, we assessed long-term trends in marginal entropy and relative entropy of annual and seasonal rainfall using the Mann–Kendall test. This will be helpful for better understanding of rainfall and streamflow characteristics, and formulating interference on climate variability, throughout the semiarid regions.

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 km^2. 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.

2.2. Rainfall and streamflow data

While other methods have been used to assess rainfall variability such as wavelet analysis (Smith et al., 2009) 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 to 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. (2001) 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: 6°24′04″ S, longitude: 38°51′44″ W, altitude: 153.4 m above sea level) and São Luiz do Curu (latitude: 3°40′12″ S, longitude: 39°14′36″ W, altitude: 38.4 m above sea level). The analyses herein are limited to two stations within the basins ranging in size from 8528 km^2 (São Luiz do Curu station) and 12,865 km^2 (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.