Extreme value theory applied to the standardized precipitation index

Gabriel Constantino Blain

Centro de Ecofisiologia e Biofísica, Instituto Agronômico, Cx. Postal 28, 13012-970, Campinas, São Paulo, Brazil. E-mail: gabriel@iac.sp.gov.br

ABSTRACT. The Standardized Precipitation Index (SPI) is a mathematical algorithm developed for detecting and characterizing precipitation departures with regard to an expected regional climate condition. Thus, this study aimed to verify the possibility of using the time-independent general extreme value distribution (GEV) for modeling the probability of occurrence of both SPI annual maxima (the maximum monthly SPI value; SPImax) and SPI annual minima (the minimum monthly SPI value; SPImim) obtained from the weather station of Campinas, State of São Paulo, Brazil (1891-2011) and to evaluate the presence of trends, temporal persistence and periodical components in these two datasets. The goodness-of-fit tests used in this study quantify the agreement between the empirical cumulative distribution and the GEV cumulative function. Our results have indicated that such parametric function can be used to assess the probability of occurrence of SPImin and SPImax values. No significant serial correlation and no trend were detected in both series. For the SPImim, the wavelet analysis has detected a dominant mode in the 4-8 year band. Future studies should focus on the development of a GEV model capable of accounting for such feature. No dominant mode was found for the annual monthly SPI maximums.

Keywords: wavelet analysis, Mann-Kendall test, SPI.

Teoria dos valores extremos aplicada à dados mensais do índice padronizado de precipitação

RESUMO. O Índice Padronizado de Precipitação (SPI) é um algoritmo matemática desenvolvido para caracterizar e detectar anomalias de precipitação pluvial em relação a uma condição climática esperada. Assim, os objetivos deste estudo foram verificar a possibilidade de utilizar a distribuição geral dos valores extremos (GEV; com parâmetros independes do tempo) para modelar a probabilidade de ocorrência dos valores mensais máximos (SPImax) e mínimos (SPImin), do referido índice de seca, obtidos a cada ano a partir da estação meteorológica de Campinas, Estado de São Paulo, Brasil (1891-2011) e, avaliar a presença de tendências, correlação serial e periodicidades em ambas as séries. Os testes de aderência utilizados quantificam a concordância entre a distribuição acumulativa empírica e a GEV. Com isso, os resultados desses procedimentos indicam que esse modelo paramétrico pode ser usado para estimar a probabilidade de ocorrência dos valores de SPImin e SPImax. Não foi observada a presença de tendências e correlações seriais em nenhuma das séries. Para o SPImim, a análise de ondaleta detectou um modo dominante na banda 4-8 anos. Recomenda-se que futuros estudos abordem o desenvolvimento de um modelo GEV capaz de incorporar em seu algoritmo essa última característica. Nenhum modo dominante foi observado para o SPImax.

Palavras-chave: análise de ondaleta, teste de Mann-Kendall, SPI.

Introduction

The Standardized Precipitation Index (SPI)[1] is a mathematical algorithm developed for detecting and characterizing precipitation departures with regard to an expected regional climate condition. According to Wu et al. (2007) the SPI is widely accepted and used throughout the world in both operational and research modes because it is (conceptually) normalized to a location and is normalized in time. In Brazil, this drought index is widely used in operational mode by governmental agricultural

institutions, such as Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), Instituto Agronômico (IAC) and Instituto Nacional de Meteorologia (INMET) in order to monitor (possible) drought conditions in different regions of the country. A review on the qualities and limitations of the SPI as well as its association with different types of drought can be found in Blain (2012a).

The aforementioned use of the SPI, associated with the fact that both positive and negative precipitation departures may cause damage to the agricultural production (LEITE et al., 2011), has allowed us to assume that evaluating the

[1] Developed by Mckee, T.B.; Doesken, N.J.; and Kleist, J.

probabilistic structure of high and low SPI values is an opportunity for increasing the knowledge related to the probability of occurrence of wet and dry events. Moreover, according to Bordi et al. (2007) and Hayes et al. (2011), the occurrence of dry or wet periods does not depend entirely on low or high precipitation totals. It also depends on a negative or a positive anomaly of these totals with respect to an expected regional climate condition. Consequently, after Bordi et al. (2007) have evaluated the nature of the tails of both precipitation and SPI distributions they indicated that this drought index is better than the precipitation for representing extreme wet and dry (monthly) events.

The statistic of extreme values is frequently used to evaluate the behavior of the M highest values of m random data (WILKS, 2011). As described in Coles (2001) and Wilks (2011), a basic result from the Theory of Extreme Values (EVT) is the so-called Extremal Types Theorem which implies that when M is stabilized with suitable sequences, it has a limiting distribution that must be one of the three types of extreme value distribution, i.e. Gumbel (Type I), Fréchet (Type II) or Weibull (Type III). These three different types can be combined into a single model; the General Extreme Value Distribution (GEV). The GEV has all the flexibility of its three particular types (BLAIN, 2011a and b; CANNON, 2010; COLES, 2001; EL ADLOUNI et al., 2007; FURIÓ; MENEU, 2010; NADARAJAH; CHOI, 2007; PUJOL et al., 2007). According to Wilks (2011) a difficulty associated with the application of the EVT is that the M values are often drawn from different distribution since they may have been generated from different physical processes. However, Wilks (2011) states that this difficulty does not invalidate this function as a candidate distribution to describe the statistics of extremes. Empirically the GEV is "[…] often an excellent choice even when the assumptions of the EVT are not met" (WILKS, 2011, p. 108). It also has to be emphasized that the Extremal Types Theorem is also applicable to distributions of extreme minima, i.e. the lowest of m records (COLES, 2001; FURIÓ; MENEU, 2010; WILKS, 2011).

The GEV is a three-parameter function such that, $\Pr\{M \leq z_t\} = GEV(z_t; \mu, \sigma, \xi)$. The Greek characters represent the time-independent parameters of this distribution. Therefore, as indicated by Coles (2001), Pujol et al. (2007), Nadarajah and Choi (2007), El Adlouni et al. (2007), Furió and Meneu, (2010), Cannon (2010) and Blain

(2011a and b) if a significant trend is found in the sample data, the use of this so-called stationary approach may no longer be valid. In other words, in the presence of a significant time-trend the probability structure of the time series does change over time (the GEV parameters vary with the time). In this sense, Blain (2011c) mentions that the evaluation of trends and other non-random components should be carried out along with the fit of a parametric distribution to a given sample data. According to authors such as El adlouni et al. (2007) and Zhang et al. (2004) nonstationarity of extreme values may be detected by evaluating the presence of trends in their time series. In this view, after Bordi et al. (2007) have applied a time-independent GEV model for assessing wet and dry periods in Sicily-Italy, they indicated that all statistical methods used in their study should be (in future efforts) adapted to the presence of climate trends.

All these statements led to the following question. Is a time-independent GEV model suitable for assessing the probability of occurrence of extreme wet and dry months observed in one of the oldest weather stations of Brazil? Thus, the aims of this study were (i) to verify the possibility of using a time-independent GEV function for modeling the probability of occurrence of SPI annual maxima (the maximum monthly SPI value; SPImax) and SPI annual minima (the minimum monthly SPI value; SPImim) obtained from the weather station of Campinas, State of São Paulo, Brazil (1891-2011) and (ii) to evaluate the presence of trends, temporal persistence and periodical components in these two sample data.

Material and methods

Annual maximums and minimums of monthly SPI data were used from the weather station of Campinas (22°54'S; 47°05'W; 669m), State of São Paulo Brazil, between 1891 and 2011. According to Blain (2009, 2012b), the rainfall monthly series observed in this location can be considered as coming from a 2-parameter gamma distribution as well as a 3-parameter Pearson Type-III distribution. The shape of these monthly probability functions vary from a strongly skewed distribution similar to those observed in arid climate (July) to a bell-shaped distribution similar to those observed in equatorial climate (January). All statistical tests used in this study were performed at the 5% level. As described in Guttman (1999) the SPI calculation starts by determining a probability density function (pdf)

capable of properly describing the long-term observed precipitation. Once a pdf is chosen, the cumulative probability [H(PRE)], associated with a given precipitation amount, are obtained from the cumulative density function [cdf (x)]. H(PRE) is then estimated from the following mixed distribution:

$$H(PRE) = (m/N) + [1 - (m/N)]cdf(x) \qquad (1)$$

where m is the number of zeros within a dataset composed of N observations. As described in Wu et al. (2007) the final step of this equiprobability transformation is based on the following rational approach proposed by Abramowitz and Stegun (1965).

$$SPI = -\left(t - \frac{co + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \text{ when } 0 < H(PRE) \le 0.5$$

$$SPI = +\left(t - \frac{co + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \text{ when } 0.5 < H(PRE) < 1 \qquad (2)$$

$$t = \sqrt{\left(\ln\left(\frac{1}{(H(PRE))^2}\right)\right)} \quad \text{when } 0 < H(PRE) \le 0.5$$

$$t = \sqrt{\left(\ln\left(\frac{1}{(1 - H(PRE))^2}\right)\right)} \quad \text{when } 0.5 < H(PRE) < 1 \qquad (3)$$

c_o = 2.515517; c_1 = 0.802853; c_2 = 0.010328; d_1= 1.432788; d_2 = 0.189269; d_3 = 0.001308
The SPI categories are shown in Table 1.

Table 1. The SPI classification system.

SPI values	Category
Greater than 2.00	Extremely wet
1.50 to 1.90	Very wet
1.00 to 1.00	Moderately wet
−0.99 to 0.99	Near normal
−1.00 to −1.00	Moderately dry
−1.50 to −1.00	Severely dry
Less than −2.00	Extremely dry

According to Guttman (1999) the Pearson type III (PE3) distribution is 'the best universal model for SPI calculations'. Furthermore, Wu et al. (2007) have stated that the confidence in SPI results, obtained from the gamma 2-parameter distribution, can be affected because this probability density function has only two free parameters. In this view, based on distinct locations of the State of São Paulo, Blain (2011d) has indicated that the SPI series obtained from the PE3 meets more frequently the normality assumption (inherent to the use of this probabilistic standardized index) than the SPI series obtained from the gamma. Therefore, the SPI values

were calculated by using the PE3 distribution. This last model was estimated as recommended by Guttman (1999) and Yue and Hashino (2007). Finally, it is worth mentioning that the SPI can be calculated for several time scales. However, as described in Bordi et al. (2007), time scales greater than the monthly scale introduces serial correlations into the SPI series. In this sense, Bordi et al. (2007) describes such component as 'an unwished property' for applying the classical EVT. Further information can be found in Bordi et al. (2007). As pointed out by Coles (2001), an extreme value analysis aims to quantify the stochastic variability of a process at unusually large levels. The so-called stationary GEV function is described as:

$$f(x) = \frac{1}{\sigma}\left[1 + \frac{\xi(SPI - \mu)}{\sigma}\right]^{1 - \frac{1}{\xi}} \exp\left\{-\left[1 + \frac{\xi(SPI - \mu)}{\sigma}\right]^{-\frac{1}{\xi}}\right\}$$

only if $1 + \frac{\xi(x - \mu)}{\sigma} > 0$ \qquad (4)

The cumulative GEV distribution is obtained by integrating the Equation 4. The parameters of the GEV distribution were estimated using the method of maximum likelihood as described in Coles (2001), Pujol et al. (2007), Nadarajah and Choi (2007) and Furió and Meneu (2010). Hereafter, these estimative will be represented by the italic font of each parameter; μ, σ, ξ and, for the sake of brevity, will be referred to as parameters. After estimating μ, σ, ξ, the return periods associated with each SPI value may also be calculated (COLES, 2001).

The chi-square test (χ^2) and the Kolmogorov-Smirnov test (KS) are frequently used to check the fit of a given dataset to a parametric distribution (WILKS, 2011). However, as pointed out by Wilks (2011) the χ^2 test operates more naturally for discrete random variable because to calculate it, the range of the data must be divided into discrete classes. On the other hand, the KS test compares the theoretical and the empirical cumulative distribution. Consequently, for continuous data, the KS test is often more powerful than the χ^2 (WILKS, 2011). Herein, the null hypothesis (H_0) of the KS test states that the data was drawn from a stationary-GEV distribution. However, if (and only if) the parameters of the theoretical distribution have not been estimated from the same data used to evaluate the fit of the parametric distribution, the original

algorithm of the KS test is applicable (STEINSKOG et al., 2007; VLČEK; HUTH, 2009; WILKS, 2011). Given that the parameters of the GEV were fitted using all available data, the KS test had to be modified. Hereafter this adapted method will be referred to as Kolmogorov-Smirnov/Lilliefors test (KS-L). The statistical simulations required for calculating the KS-L test were based on the procedure called 'non-uniform random number generation by inversion'. It were generated $Ns = 10000$ synthetic data samples. More information about the KS-L can be found in Wilks (2011).

Regarding extreme-value statistics, the Anderson-Darling test (AD) (ANDERSON; DARLING, 1952) is based on the sum of the squares of the differences between theoretical and empirical distribution with a weight function [$\Psi(SPI)$] which emphasizes the discrepancies in both tails (SHIN et al., 2012). This last feature cannot be observed in the KS-L algorithm. The AD test is based on a class of quadratic function, such that

$$Qn = N \int_{-\infty}^{\infty} [F(SPI) - G(SPI)]^2 \, \Psi(SPI) dG(SPI) \quad (5)$$

where N is the length of the series, G(SPI) is the cumulative distribution function, and F(SPI) is the empirical distribution.

As described in Shin et al. (2012) when [$\Psi(SPI)$] = 1, Qn becomes equal to the Cramer von Mises test. When $\Psi(SPI) = \{G(SPI)[1-G(SPI)]\}^{-1}$; Qn becomes equal to the AD test. This last form of calculating $\Psi(SPI)$ (and consequently Qn) usually leads to a more powerful test by emphasizing the tails differences (SHIN et al., 2012). According to Haddad and Rahman (2011) the AD test is better than other tests, such as Akaike Information Criterion and Bayesian Information Criterion, in recognizing the parent distribution when this last one is a 3-parameter function. As can be noted from equation 5 the AD gives equal weights to both upper and lower tails of the distribution (SHIN et al., 2012). However, by following Shin et al. (2012) we may assume that studies addressing drought events are basically interested in the lower tail of the (SPI) distributions. By way of analogy, studies addressing wet events are fundamentally interested in the upper tail of the (SPI) distributions. In both cases, the use of a $\Psi(SPI)$ that gives emphasis to discrepancies at either upper or lower tails of the distributions becomes an interesting choice. On such a background, we describe the modified Anderson-Darling test (AHMAD et al, 1988), in which, for emphasis on the upper tail, $\Psi(SPI)$ is set to be equal to $[1-G(SPI)]^{-1}$ (Equation 6). For emphasis on the lower tail, $\Psi(SPI)$ is set to be equal to $[G(SPI)]^{-1}$ (Equation 7).

$$AU = N \int_{-\infty}^{\infty} \frac{[F(SPI) - G(SPI)]^2}{1 - G(SPI)} \Psi(SPI) dG(SPI) \quad (6)$$

$$AL = N \int_{-\infty}^{\infty} \frac{[F(SPI) - G(SPI)]^2}{G(SPI)} \Psi(SPI) dG(SPI) \quad (7)$$

As described in Shin et al. (2012) AU + AL = AD. As for the KS-L, the statistical simulations required for assessing the significance of AU, AL, and AD were obtained from the 'non-uniform random number generation by inversion' (Ns = 10000).

Temporal persistence, trends and periodical components

A cdf summary may conveniently represent the probabilistic information obtained from a time series (MAIA et al., 2007). However, this convenient way is only appropriated in the presence of a non-significant serial correlation. If the sample data exhibits significant auto-correlations such summary leads to a loss of information (MAIA et al., 2007). Therefore, the run test and the Durbin-Watson test (DW) were applied to verify if the time series can be considered as being free from temporal persistence (p-values above 0.05 indicate that the time series is free from significant serial correlations). For the same purpose, the significance of the lag-1 auto-correlation coefficient (r_1) was also evaluated. Further information regarding the r_1 calculation can be found in Wilks (2011).

The Mann-Kendall (MK) test (KENDALL; STUART, 1967) was used for evaluating the presence of trends in both SPImin and SPImax series. The null hypothesis (H_0) associated with this widely used trend test assumes that the data is independently and identically distributed (iid). Under practical applications the rejection of such H_0 is often taken as an evidence of the presence of trend in a given time series. By following Chandler and Scott (2011) we may assume that this last assumption is reasonable since in idd data no trend is present. However, the

evaluation of the presence of serial correlation is a fundamental step for avoiding a false rejection of H_0 (BLAIN; PIRES, 2011; BURN et al., 2004; SANSIGOLO; KAYANO, 2010) (among many others). The significance of each MK outcome was obtained from a bootstrap procedure considering Ns = 10000 trend-free synthetic sample data. Further information related to the MK calculation algorithm used in this study can be found in Burn et al. (2004). Although the significance of each MK statistics can also be obtained from the normal cumulative distribution function (BLAIN, 2011a; DUFEK; AMBRIZZI, 2008) (among many others) the bootstrap approach can be applied with any statistical test for trend detection (BURN; ELNUR, 2002). According to Yue and Pilon (2004) both approaches (bootstrap-based and normality assumption) have the same power. By following Torrence and Compo (1998), we applied the wavelet analysis to decompose both SPImax and SPImim series into time-frequency space. This spectral analysis has allowed (i) to observe the variance peaks in the frequency domain and (ii) to verify how those peaks vary in time. Detailed explanation of the wavelet technique, including its statistical significance testing, can be found in Torrence and Compo (1998). The wavelet function used in the present study was the Morlet wavelet (BEECHAM; CHOWDHURY, 2010; BLAIN 2012b; PEZZI; KAYANO 2009). According to Souza-Echer et al. (2008) the complex Morlet wavelet is the most suitable to evaluate variations of periodicities in geophysical signals. The computational algorithm used for calculating this method is available at http://paos.colorado.edu/research/wavelets.

Results and discussion

The KS-L, AD, AU and AL tests quantify the agreement between the empirical cumulative distribution and the theoretical cumulative function (SHIN et al., 2012; WILKS, 2011). Therefore, these tests provide empirical evidences for accepting or rejecting the hypothesis that a given theoretical distribution may be used to assess the probability of occurrence of a given variable. In this view, all the goodness-of-fit tests have indicated that the time independent GEV model may be used to evaluate the probability of occurrence of the SPImax as well as the SPImim obtained from the weather station of Campinas. The p-values associated with each outcome were far from the 5% critical level.

Regarding dry events, the results obtained from the tests were: KS-L = 0.0432 (p-value = 0.94), AD = 0.381 (p-value = 0.87), AL = 0.159 (p-value = 0.90), and AU = 0.221 (p-value = 0.77). Regarding wet events, the obtained outcomes were: KS-L=0.027 (p-value > 0.99), AD = 0.098 (p-value > 0.99), AL = 0.048 (p-value > 0.99), and AU = 0.050 (p-value > 0.99). Thus, we have enough empirical evidences to assume that both SPImax and SPImim observed during 1891-2011 in Campinas were drawn from a GEV distribution with the following parameters: μ_{SPImim} =-1.32 [-1.41:-1.21], σ_{SPImim} = 0.50 [0.43:0.57], ξ_{SPImim} = -0.12 [-0.21:-0.02] (Dry events); μ_{SPImax} = -1.37 [1.27:1.47], σ_{SPImax} = 0.52 [0.45:0.59], ξ_{SPImax} = -0.18 [-0.30:-0.07] (Wet events); [.] represents the 95% confidence interval.

As described in several studies such as Coles (2001), and Wilks (2011), the negative values of ξ have defined the third special case of the GEV distribution known as Weibull. Since this last parameter determines the rate of tail decay, such negative values have defined short-tailed (SPI) distributions (COLES, 2001).

In addition, as described in Delgado et al. (2010), while μ defines the position of the GEV function with regard to the origin, σ defines the spread of the distribution. In this view, the confidence intervals associated with the estimates of σ (dry and wet events) and ξ (dry and wet events) overlap. Thus, the spread and the tail behavior of both SPImax and SPI min distributions were similar to each other. Table 2 lists a practical application of these stationary GEV models.

Aiming to reveal the effects of arid climate or those with a distinct dry season on the SPI values, Wu et al. (2007) used a mathematical algorithm (which may be seen as a normality test) in order to investigate whether the SPI values, obtained from different precipitation regimes across the United States, represent drought and flood events in a similar (probabilistic) way. Based on the same mathematical algorithm, Blain (2012b) identified that the SPI monthly values, obtained from the weather station of Campinas and computed from the PE3 function, can be seen as coming from a normal distribution with zero mean and unit variance. Therefore, from statistical inference theory we may assume that the probability of occurring a SPI value equal to or lower than -2 in a given month is (only) 0.02.

Table 2. Probabilities of occurrence of annual maximums and minimums of monthly-SPI values. Campinas, State of São Paulo, Brazil. The associated return period (RP) are also shown.

SPI	Pr{Z ≤ SPI}	RP (1/Pr{Z ≤ SPI})	SPI	Pr{Z > SPI}	RP (1/Pr{Z > SPI})
-1.0	0.841	1	1.0	0.858	1
-1.5	0.500	2	1.5	0.537	2
-2.0	0.201	5	2.0	0.221	5
-2.5	0.061	16	2.5	0.058	17

In spite of this relative low probability, the probability of having at least one extreme dry month (SPI ≤ -2.0; Table 1) within a given year rises to 0.20 (Table 2; weather station of Campinas). By analogy, while the return period associated with a monthly-SPI value equal to -2.0 is 50 years, the return period associated with the probability of having at least one extreme dry month within a given year is (only) 5 years. Similar conclusions can be taken for any other SPI value. From Table 2, the return period associated with SPI values between -2.5 and -1.0 are similar to those related to SPI values between 1.0 and 2.5. This feature agrees with Bordi et al. (2007). According to these authors, both wet and dry SPI values have similar extreme behavior.

Temporal persistence, trends, and periodical components

The results obtained through the run and DW tests indicate no significant serial correlation in the analyzed datasets. Regarding the SPImin, the p-values obtained from both tests are, respectively, 0.78 and 0.56. Regarding the SPImax, the p-values obtained from both tests are, respectively, 0.31 and 0.68.

These results agree with those obtained from the equation 7. The r_1 values (-0.0832; SPImin and 0.013; SPImax) remained within the 5% significance interval. Thus, following Maia et al. (2007), we accepted that a cdf summary of both SPImin and SPImax series will result in loss of no significant information. The MK test indicated no significant trend in the SPImin and SPImax series (Figure 1).

By following Coles (2001), Pujol et al. (2007), Nadarajah and Choi (2007), Furió and Meneu (2010), Cannon (2010) and Blain (2011a, b and c) we may interpret this lack of significant trends in the SPI series as supporting the use of the time-independent approach adopted in this study.
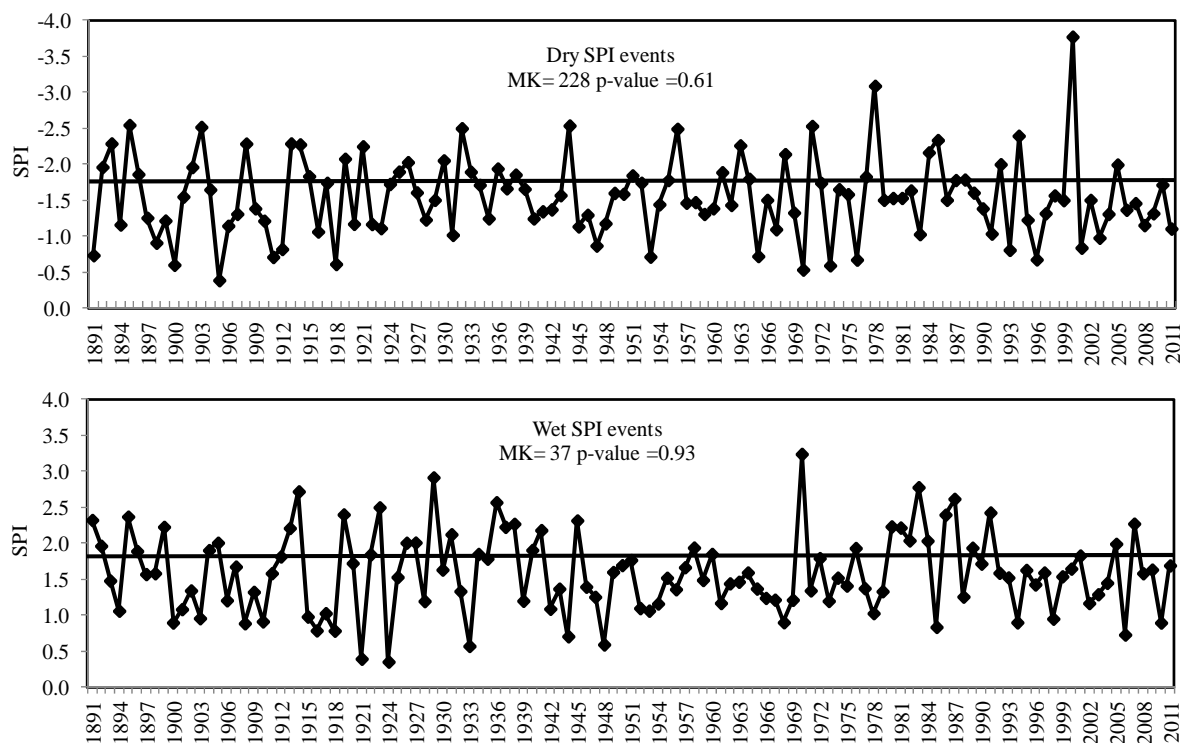

Figure 1. Annual maximum monthly SPI (wet events) and annual minimum monthly SPI (dry events) obtained from the weather station of Campinas, State of São Paulo, Brazil. The Mann-Kendall trend test (MK) and the associated p-value are also shown. Horizontal lines represent the mean of each series.

As described by Beecham and Chowdhury (2010) the dashed line in the Global Wavelet Power (GWP; Figure 2b) represents the 95% confidence level. A given peak of variance, at any frequency level (f) that crosses this dashed line indicates a statistical significance for such (T = 1/f) period (BEECHAM; CHOWDHURY, 2010). Therefore, Figure 2b shows a significant variance peak with a frequency f that may be described, approximately, as $f = 0.17yr^{-1}$. In other words, the Figure 2b indicates that the dominant mode of temporal variability of the SPImin is within the 4-8 year band. Considering the time-frequency domain, the wavelet power spectrum (Figure 2a) showed intermittent concentration of energy within this band.

For instance, there are appreciable powers from 1900-1920 and from 1995-2002 (the area under the cone influence is not considered). Figure 2a also presented other (non-significant) variance peaks (1940-1960 and 1975-1985). Therefore, we may indicate that the Figure 2 provided evidences for accepting that the SPImin series has a dominant mode in the 4-8 year band. Considering the 2-4 year band, Figure 2a also showed other two periods of enhanced variance, 1965-1980 and 1995-2002. Investigating the physical phenomena that lead to such dominant mode is beyond the scope of this study.

However, it is worth mentioning that after Torrence and Compo (1998) have applied the wavelet analysis to Niño 3 Sea Surface Temperature, they verified a concentration of wavelet power within the El-Niño Southern Oscillation band (2-8 years). Thus, future studies should investigate a possible coherence between these two signals (SPImin and Niño 3 sea surface temperature). Finally, it has to be emphasized that the presence of a dominant mode indicates the need to consider such feature when the probability of occurrence of the SPImin values is being assessed. In other words, by way of analogy with Bordi et al. (2007), we recommend that the GEV model used herein, should be (in further studies) adapted to the presence of such periodical components.

The lack of significant variance peaks in the spectral analysis, carried out in the frequency domain (Figure 3b; GWP), supported the hypothesis of no significant periodical component in the SPImax signal. In this sense, the wavelet power spectrum (time-frequency domain) presented a concentration of energy only within localized regions of Figure 3a (WPS). Considering the 4-8 year scale, it was observed significant power only between 1900 and 1920. For time scales shorter than 5 year, there are two significant variance peaks (1918-1923 and 1981-1983). Thus, the wavelet analysis showed the presence of no dominant mode in the SPImax temporal variability.

Finally, the results found in this study allowed stating that future studies should verify the possibility of using the GEV function for modeling the probability of occurrence of both SPImax and SPImim obtained from other regions of interest. In this view, the procedures and the results described herein may be seen as a reference.
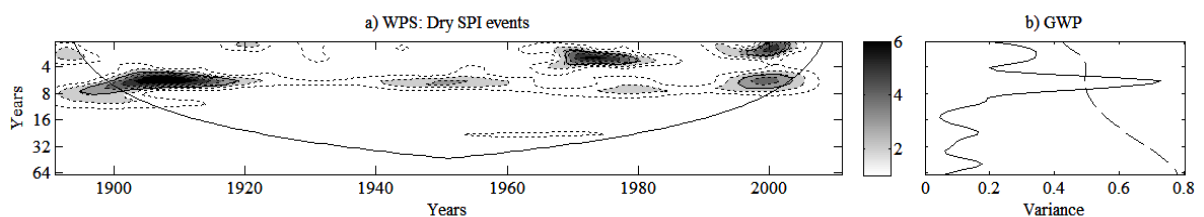

Figure 2. (a) Wavelet power spectrum (WPS) of the annual minimum monthly SPI. Campinas, State of São Paulo, Brazil (normalized by 1/0.311); (b) Global Power Spectrum (GWP - in variance units).
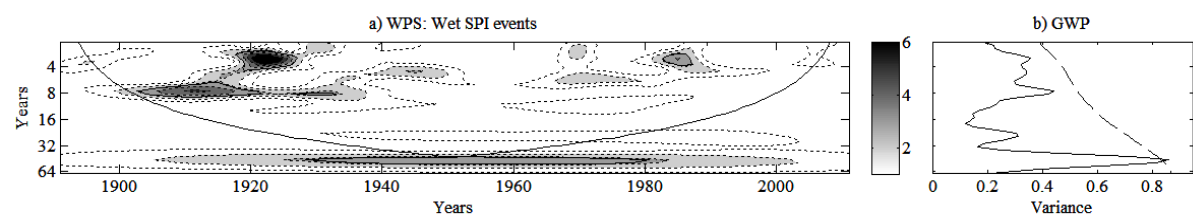

Figure 3. (a) Wavelet power spectrum (WPS) of the annual maximum monthly SPI. Campinas, State of São Paulo, Brazil (normalized by 1/0.305²); (b) Global Power Spectrum (GWP - in variance units).

Conclusion

The time-independent GEV function can be used to assess the probability of occurrence of SPI annual maxima and minima (maximum and minimum monthly SPI values) obtained from Campinas, State of São Paulo, Brazil.

For the SPI annual minima, the spectral analysis has detected a dominant mode in the 4-8 year band. This variability was not evenly distributed over the analyzed period. Therefore, the evaluation of a possible relationship between this dominant mode and a non-stationary geophysical process should be addressed in future studies. No dominant mode was detected for the SPI annual maxima.

References

ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of mathematical function**. Dover: Dover Publications, 1965.

AHMAD, M. I.; SINCLAIR, C. D.; SPURR, B. D. Assessment of flood frequency models using empirical distribution function. **Water Resource Research**, v. 24, n. 8, p. 1323-1328, 1988.

ANDERSON, T. W.; DARLING D. A. Asymptotic theory of certain 'goodness-of-fit' criteria based on stochastic processes. **The Annals of Mathematical Statistics**, v. 23, n. 2, p. 193-212, 1952.

BEECHAM, S.; CHOWDHURY, R. K. Temporal characteristics and variability of point rainfall: a statistical and wavelet analysis. **International Journal of Climatology**, v. 30, n. 3, p. 458-473, 2010.

BLAIN, G. C. Considerações estatísticas relativas à oito séries de precipitação pluvial da Secretaria de Agricultura e Abastecimento do Estado de São Paulo. **Revista Brasileira de Meteorologia**, v. 24, n. 1, p. 12-23, 2009.

BLAIN, G. C. Cento e vinte anos de totais extremos de precipitação pluvial máxima diária em Campinas, Estado de São Paulo: análises estatísticas. **Bragantia**, v. 70, n. 3, p. 722-728, 2011a.

BLAIN, G. C. Modeling extreme minimum air temperature series under climate change conditions. **Ciência Rural**, v. 41, n. 11, p. 1877-1883, 2011b.

BLAIN, G. C. Incorporating climate trends in the stochastic modeling of extreme minimum air temperature series of Campinas, state of São Paulo, Brazil. **Bragantia**, v. 70, n. 4, p. 234-245, 2011c.

BLAIN, G. C. Standardized Precipitation Index based on Pearson type III distribution. **Revista Brasileira de Meteorologia**, v. 26, n. 2, p. 167-180, 2011d.

BLAIN, G. C. Revisiting the probabilistic definition of drought: strengths, limitations and an agrometeorological adaptation. **Bragantia**, v. 71, n. 1, p. 132-141, 2012a.

BLAIN, G. C. Monthly values of the standardized precipitation index in the State of São Paulo, Brazil: trends and spectral features under the normality assumption. **Bragantia**, v. 71, n. 1, p. 460-470, 2012b.

BLAIN, G. C.; PIRES, R. C. M. Variabilidade temporal da evapotranspiração real e da razão entre evapotranspiração real e potencial em Campinas, Estado de São Paulo. **Bragantia**, v. 70, n. 2, p. 460-470, 2011.

BORDI, I.; FRAEDRICH, K.; PETITTA, M.; SUTERA, A.; Extreme value analysis of wet and dry periods in Sicily. **Theoretical and Applied Climatology**, v. 87, n. 1-4, p. 61-71, 2007.

BURN, D. H.; ELNUR, A. H. Detection of hidrological trends and variability. **Journal of Hydrology**, v. 255, n. 255, p. 107-122, 2002.

BURN, D. H.; CUNDERLIK, J.; PIETRONIRO, A. Hydrological trends and variability in the Liard River basin. **Hydrological Sciences**, v. 49, n. 1, p. 53-67, 2004.

CANNON, A. J. A flexible nonlinear modeling framework for nonstationary generalized extreme value analysis in hydroclimatology. **Hydrological Process**, v. 24, n. 1, p. 673-685, 2010.

CHANDLER, R. E.; SCOTT, M. E. **Statistical methods for trend detection and analysis in the environmental analysis**. Chichester: John Wiley & Sons, 2011.

COLES, S. **An introduction to statistical modeling of extreme value**. London: Springer, 2001.

DELGADO, J. M.; APEL, H.; MERZ, B. Flood trends and variability in the Mekong river. **Hydrology and Earth System Sciences**, v. 14, n. 1, p. 407-418, 2010.

DUFEK, A. S.; AMBRIZZI, T. Precipitation variability in Sao Paulo State, Brazil. **Theoretical and Applied Climatology**, v. 93, n. 3-4, p. 167-178, 2008.

EL ADLOUNI, S.; OUARDA, T. B. M. J.; ZHANG, X.; ROY, R.; BOBÉE, B. Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. **Water Resources Research**, v. 43, n. 3410, p. 1-13, 2007.

FURIÓ, D.; MENEU, V. Analysis of extreme temperatures for four sites across Peninsular Spain. **Theoretical Applied Climatology**, v. 104, n. 1, p. 83-99, 2010.

GUTTMAN, G. B. Accepting the Standardized Precipitation Index: a calculation algorithm index. **Journal of the American Water Resources**, v. 35, n. 2, p. 311-322, 1999.

HADDAD, K., RAHMAN, A. Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for Tasmania in Australia. Stochastic **Environmental Research and Risk Assessment**, v. 25, n. 3, p. 415-428, 2011.

HAYES, M. J.; SVOBODA, M. D.; WALL, N.; WIDHALM, M. The Lincoln declaration on drought indices - universal meteorological drought index recommended. **Bulletin of the American Meteorological Society**, v. 92, n. 4, p. 485-488, 2011.

KENDALL, M. A.; STUART, A. **The advanced theory of statistics**. 2nd ed. Londres: Charles Griffin, 1967.

LEITE, M. L.; ADACHESKI, P. A.; VIRGENS FILHO, J. S. Análise da frequência e da intensidade das chuvas em Ponta Grossa, Estado do Paraná, no período entre 1954 e 2001. **Acta Scientiarum. Agronomy**, v. 33, n. 1, p. 54-57, 2011.

MAIA, A. H. N.; MEINKE, H.; LENNOX, S.; STONE, R. C. Inferential, non-parametric statistics to assess quality of probabilistic forecast systems. **Monthly Weather Review**, v. 135, n. 2, p. 351-362, 2007.

NADARAJAH, S.; CHOI, D. Maximum daily rainnfall in South Korea. **Journal of Earth System Science**, v. 116, n. 4, p. 311-320, 2007.

PEZZI, L. P.; KAYANO, M. T. An analysis of the seasonal precipitation forecasts in South America using wavelets. **International Journal of Climatology**, v. 29, n. 11, p. 1560-1573, 2009.

PUJOL, N.; NEPPEL, L.; SABATIER, R. Regional tests for trend detection in maximum precipitation series in the French Mediterranean region. **Hydrological Sciences Journal**, v. 52, n. 5, p. 952-973, 2007.

SANSIGOLO, C. A.; KAYANO, M. T. Trends of seasonal maximum and minimum temperatures and precipitation in Southern Brazil for the 1913–2006 period. **Theoretical and Applied Climatology**, v. 101, n. 1-2, p. 209-216, 2010.

SHIN, H.; JUNG, Y.; JEONG, C. Assessment of modified Anderson–Darling test statistics for the generalized extreme value and generalized logistic distributions **Stochastic Environmental Research Risk Assessment**, v. 26, n. 1, p. 105-114, 2012.

SOUZA-ECHER, M. P.; ECHER, E.; NORDEMANN, D. J.; RIGOZO, N. R.; PRESTES, A. Wavelet analysis of a centennial (1895-1994) southern Brazil rainfall series (Pelotas, 31°46'19"S 52°20'33"W). **Climate Change**, v. 87, n. 3, p. 489-497, 2008.

STEINSKOG, D. J.; TJØSTHEIM, D. B.; KVAMSTØ, N. G. A cautionary note on the use of the Kolmogorov–Smirnov test for normality. **Monthly Weather Review**, v. 135, n. 3, p. 1151-1157, 2007.

TORRENCE, C.; COMPO, G. P. Practical guide to Wavelet analysis. **Bulletins of American Meteorological Society**, v. 79, n. 1, p. 61-78, 1998.

VLČEK, O.; HUTH R. Is daily precipitation Gamma-distributed? Adverse effects of an incorrect application of the Kolmogorov-Smirnov test. **Atmospheric Research**, v. 93, n. 4, p. 759-766, 2009.

WILKS, D. S. **Statistical methods in the atmospheric sciences**. San Diego: Academic Press, 2011.

WU, H.; SVOBODA, M. D.; HAYES, M. J.; WILHITE, D. A.; WEN, F. Appropriate application of the Standardized Precipitation Index in arid locations and dry seasons. **International Journal of Climatology**, v. 27, n. 1, p. 65-79, 2007.

YUE, S.; HASHINO, M. Probability distribution of annual, seasonal and monthly precipitation in Japan. **Hydrological Sciences Journal**, v. 52, n. 5, p. 863-877, 2007.

YUE, S.; PILON, P. A comparison of the power of the t test, Mann-Kendall and bootstrap tests for trend detection. **Hydrological Sciences Journal**, v. 49, n. 49, p. 53-37, 2004.

ZHANG, X.; ZWIERS, F.; W.; LI, G. Monte Carlo experiments on the detection of trends in extreme values. **Journal of Climate**, v. 17, n. 10, p.1945-1952, 2004.