Drought analysis in southern Paraguay, Brazil and northern Argentina: regionalization, occurrence rate and rainfall thresholds

Maria Manuela Portela, João Filipe dos Santos, Artur Tiago Silva, Julián Baez Benitez, Carlos Frank and José Miguel Reichert

ABSTRACT

The objective of the present study is to characterize the drought occurrences in a region comprising Paraguay, southern Brazil and northeastern Argentina. To recognize the drought occurrences the standardized precipitation index at the time-scales of 3 and 6 months was applied to the rainfall records from 1961 to 2011 at 51 rain gauges located in that region. After a drought regionalization using principal component analysis, a new approach, the Kernel occurrence rate estimation method coupled with bootstrap confidence band was used to quantify yearly drought occurrence rates. The study also includes the results of an additional and new approach based on rainfall threshold surfaces aimed at recognizing and monitoring the drought occurrence at the early stages of their development. For both time-scales, the study allowed identification of some spatial homogeneous regions regarding the severe droughts. In some of those regions, trends in the severe drought frequency occurrence were identified. The rainfall threshold surfaces, besides providing an adequate interpretation of the meaning of the standard precipitation index, can be quite easily and reliably utilized to identify the drought episodes.

Key words | drought, Kernel occurrence rate estimator, principal component analysis, rainfall threshold surfaces, standardized precipitation index

Maria Manuela Portela
Artur Tiago Silva
Universidade de Lisboa,
Instituto Superior Técnico,
Lisbon, Portugal

João Filipe dos Santos (corresponding author)
Instituto Politécnico de Beja,
Escola Superior de Tecnologia e Gestão de Beja,
Beja, Portugal
E-mail: *Joaof.santos@ipbeja.pt*

Julián Baez Benitez
Universidad Católica 'Nuestra Señora de la Asunción',
Facultad de Ciencais y Tecnología,
Asunción, Paraguay

Carlos Frank
Universidad Tecnológica Nacional,
Facultad Regional Bahí-a Blanca,
Bahía Blanca, Argentina

José Miguel Reichert
Universidade Federal de Santa Maria,
Centro de Ciências Rurais,
Santa Maria, Brazil

INTRODUCTION

Droughts are generally associated with the persistence of low rainfall, soil moisture and water availability relative to the normal levels in a designated area. Although there is no universally accepted definition for drought, Tallaksen & Van Lanen (2004) defined it as 'a sustained and regionally extensive occurrence of below average natural water availability'. Different from other extreme events, like floods and earthquakes, droughts remain a less visible natural risk, whose impacts are not systematically recorded. Droughts are among the most complex and least understood natural hazards, affecting more population than any other. They are also recurrent hazards particularly in areas with pronounced natural climate temporal variability, as those under analysis.

Some authors have studied drought occurrences in several countries. Recently Lee *et al.* (2012) analyzed the spatiotemporal characteristics of drought occurrences in Japan for the period from 1902 to 2009 using an effective drought index. With hierarchical cluster analysis applied to drought characteristics data (such as duration, severity, and onset and end dates) available at 50 monitoring stations, drought regions were identified and drought occurrences evaluated.

Using future climate scenarios, Sheffield & Wood (2008) analyzed changes in drought occurrence using soil moisture data. The models showed decreases in soil moisture globally for all scenarios with a corresponding doubling of the spatial extent of severe soil moisture deficits and frequency of

doi: 10.2166/nh.2014.074

short-term (4–6-month duration) droughts from the mid-20th to the end of the 21st centuries. Long-term droughts become three times more common. Regionally, the Mediterranean, west African, central Asian and central American regions show large increases most notably for long-term frequencies as do mid-latitude North American regions but with larger variation between scenarios.

In Argentina some related studies have also been undertaken, namely the work of Capriolo & Scarpati (2012), who used the soil water balance obtained by the evapotranspiration formula of Penman–Monteith to consider soil water deficit and surplus as triggers of extreme hydrologic events. The authors considered annual threshold values of 200 mm of soil water deficit and 300 mm of soil water surplus for drought and flood recognition, respectively. Using the Mann–Kendall statistical test, the results have shown significant trends at level 0.1 for drought in two periods, 1 of 20 years (1991–2010) and the other of 10 years (2001–2010). Ravelo (2000) analyzed droughts for the period of 1931–1999 in the plains region of Argentina. Meteorological drought indices were used in a time and space analysis to establish drought intensity, frequency, probability distribution and probabilities of occurrence of given drought intensities. More extensively Barrucand *et al.* (2007) studied the frequency and spatial distribution of droughts in different regions of Argentina during the 20th century. The behavior of the mean monthly atmospheric circulation associated with dry conditions in the Pampas during the second half of the century was analyzed.

Within the study, Paraguay is particularly affected by droughts. The potential impacts of such hazards on the economy of the country can be significant, especially taking into account that Paraguay is the sixth largest producer of soybean in the world (Masuda & Goldsmith 2009). As a consequence of one of the most severe drought periods, from November 2008 to March 2009, the gross domestic product fell by 4.2% in the first trimester of 2009, and the yield of soybean suffered a reduction of 39% (Inter-American Development Bank, 2014). In Argentina, the impact of the event was even worse.

In southern Brazil, over an area including the states of Paraná, Santa Catarina and Rio Grande do Sul some historical droughts have also been investigated (Mattos *et al.* 2012). The authors using the standard precipitation index (SPI) have identified nine remarkable droughts since 1961 the most severe being the one that occurred in 1985, with 94% of the study area under extreme drought, and the longest occurring in 2006 with 12 consecutive months under drought. They also found forcing links between Tropical Pacific and Indian Ocean sea surface temperatures (SSTs) and the SPI.

Impact assessment of a specific drought requires knowing its causes and the spatial and temporal distribution of the rainfall anomalies. Grimm *et al.* (2000) and Grimm (2004) analyzed the influence of El Niño Southern Oscillation warm (El Niño) and cold (La Niña) phases in the rainfall patterns of the southeastern region of South America, providing a comprehensive view of the anomalies of rainfall and atmospheric circulation associated with both ocean–atmosphere phase events. The La Niña phase coincides with a reduction in rainfall in northern Argentina, southern Brazil and southeastern Paraguay. Fraisse *et al.* (2008) analyzed soybean yields in Paraguay and rainfall amounts during different phases of crop phenological development and found significant rainfall reductions, especially between planting and blooming during La Niña years.

The perception of the meaning of drought and its impacts varies significantly (Vogt & Somma 2000) as, in fact, drought does not have a precise and universally accepted definition. Nevertheless, it should be stated that there is a consensus regarding the following different types of drought: meteorological, agricultural, hydrological and the socioeconomic (Wilhite & Glantz 1985). Such types of drought can also be defined in direct connection with the SPI, developed by McKee *et al.* (1993) aiming at quantifying the rainfall deficit at multiple time-scales.

Meteorological drought is caused by a rainfall deficit over an extended period of time. This deficit may be accumulated and expressed relative to a climate norm and to the duration of the dry period (Lloyd-Hughes 2002). Definitions of meteorological drought must be considered as region specific since the atmospheric conditions that result in deficiencies of rainfall are highly variable from region to region (Wilhite 1994). The SPI is linked to this drought type when calculated at an approximately 1–3-month timescale (Hayes *et al.* 1999).

The water soil deficiency is usually connected to agricultural drought and is caused by a deficit of fresh water

relative to evapotranspiration losses. A drought exists when the water availability at the root-zone is insufficient to sustain crops and pasture between rainfall events (Tate & Gustard 2000). For agricultural drought, Sims *et al.* (2002) reported a strong relationship between SPI over short time-scales (estimated 3–6 months) and temporal variations of soil moisture.

A hydrological drought results directly from reduced rainfall, which originates from reduced surface runoff and, indirectly, from reduced groundwater discharge to the river channel. Key indicators are reduced river flows and low water levels in lakes and reservoirs. According to Lloyd-Hughes (2002), hydrological droughts are the most visible and important in terms of human perception. According to some authors, the SPI at a 12-month time-scale could be considered a hydrological drought index, having been tested for monitoring surface water resources, e.g. river flows and water levels in lakes (Hayes *et al.* 1999; Szalai & Szinell 2000).

At longer time-scales of the SPI (24 or 36 months), droughts last longer, but are less frequent. They are used to monitor the impact of droughts on aquifers, which are systems that respond more slowly to changing conditions (Changnon & Easterling 1989).

In the applications carried out, the time-scales considered for the SPI calculation were 3 and 6 months, since they provide a useful monitoring tool for agricultural drought assessment, which is particularly convenient due to the importance of the economic sector in the region under study.

The aim of this work was to investigate some of the drought characteristics in a study region, which comprises Paraguay, southern Brazil and northeastern Argentina in an attempt to understand the historical and recent climatic variability. The analysis utilized 51 years of rainfall data (1961–2011) in 51 rain gauges fairly distributed over the region. To recognize the drought occurrences, the SPI at 3 and 6 months was computed based on the previous monthly rainfall data. The novelty of this work, especially in regard to the study area, is (i) the identification of spatial patterns of droughts, (ii) the characterization of the yearly drought occurrence rates using the Kernel occurrence rate estimation (KORE) method coupled with bootstrap confidence band and (iii) the establishment of surfaces of rainfall thresholds for drought recognition, obtained by inverting the SPI values that represent drought thresholds back to the rainfall field (Santos *et al.* 2013), thus facilitating an adequate interpretation of the meaning of such an index and quite easily and reliably identifying the drought episodes.

This paper is organized into five sections, two of them with subsections. With the background provided by this introductory section, the models applied are then described which comprises a briefly description of the drought index used (the SPI) and of the models applied in its characterization either in spatial terms (the principal component analysis (PCA)) or in terms of the occurrence rate of extreme droughts (the KORE estimation method). The study region and rainfall data are then presented followed by the results, which besides the analysis related to the identification of homogenous regions and to the changes in the occurrence rate of the droughts, includes surfaces of cumulative rainfall thresholds for drought recognition. Finally, conclusions are drawn and future research scenarios are proposed.

METHODS

Recognition of the drought occurrences

The analysis of the temporal variability of droughts was assessed via the SPI, one of the most popular and common drought indices (Vicente-Serrano 2006; Santos *et al.* 2010).

The SPI, originally developed by McKee *et al.* (1993), remaps the rainfall records into a standardized probability distribution function so that an index of zero indicates the median rainfall amount, while a negative index stands for drought conditions and a positive index for wet conditions (Santos *et al.* 2011). A comprehensive description of the calculation algorithm and of the advantages of the SPI index can be found in the literature (Edwards & McKee 1997; Guttman 1998, 1999; Hayes *et al.* 1999; Lloyd-Hughes & Saunders 2002; Santos & Portela 2010).

As summarized by Santos *et al.* (2010), there are several advantages to the SPI, namely: (i) its flexibility, as it can be applied at various time-scales; (ii) there is less complexity involved in its implementation, relative to other drought indices; (iii) it is adaptable to other hydroclimatic variables

besides rainfall (Santos & Portela 2010); (iv) its suitability for spatial analysis, allowing comparison between sites in a given region as it is a normalized index.

Originally the computation of the SPI index utilized a Gamma distribution function applied to the monthly rainfall series, for SPI1, and to cumulative rainfall series, for the other time-scales (McKee *et al.* 1993). Later, several authors tested different statistical distributions based on different time-scales and concluded that, compared to the Gamma distribution, the Pearson type III distribution ensured the best fit due to its higher flexibility given by its three parameters (Guttman 1999; Ntale & Gan 2003; Vicente-Serrano 2006). The Pearson III probability distribution function is given by

$$f(x) = \frac{1}{\alpha\Gamma(\beta)}\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1}\exp\left(-\frac{x-\gamma}{\alpha}\right) \quad (1)$$

where γ, α and β are the location, scale and shape parameters, respectively. The parameters of the previous distribution were estimated using the L-moments method.

The value of SPI attributed to each rainfall amount is the z-standard normal associated to the probability of non-exceedance of that rainfall, according to the Pearson type III distribution, as represented in Figure 1.

The relationship among the SPI values, the probabilities of non-exceedance and the drought categories adopted in the study presented are shown in Table 1.

Table 1 | Drought category thresholds given by the standard precipitation index and the respective non-exceedance probabilities (Agnew 2000)

Probability	SPI	Drought category
0.05	>1.65	Extremely wet
0.10	>1.28	Severely wet
0.20	>0.84	Moderately wet
0.60	> −0.84 and <0.84	Normal
0.20	< −0.84	Moderate drought
0.10	< −1.28	Severe drought
0.05	< −1.65	Extreme drought

The SPI calculated at different time-scales can be linked to different drought definitions, which refer to one or more components of the hydrological cycle (Rossi 2003), as previously specified.

Drought spatial patterns

One of the main objectives of the present study was to identify spatial patterns of droughts at different time-scales on the study area based on the rainfall records. To do that PCA was used.

PCA is a regionalization technic that can be used to identify homogenous groups of variables that experienced similar drought (or wet) conditions during a study period (Bonaccorso *et al.* 2003; Vicente-Serrano *et al.* 2004;

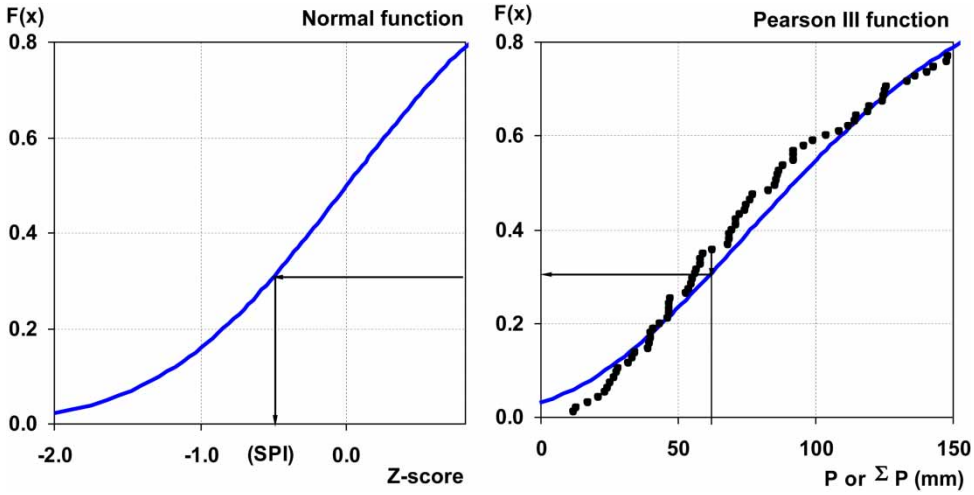

Figure 1 | Schematic representation of the SPI calculation procedure – reproduced from Santos *et al.* (2013).

Santos *et al.* 2010) and for which can be ascribed a physical meaning (Ehrendorfer 1987).

Some authors, such as Lins (1985), Tipping & Bishop (1999), Jolliffe (2002), Kahya *et al.* (2008a, b), Westra *et al.* (2007) or Singh *et al.* (2009), define the PCA method as a technique that allows decomposing the multisite data set of a given variable (e.g. the SPI field) into univariate representations of that variable. So the values of these new representations of the variable can be interpreted geometrically as the projections of its observations onto the principal components (PCs) (Abdi & Williams 2010). By that way, the original intercorrelated variables can be reduced to a small number of new linearly uncorrelated ones that explain most of the total variance (Rencher 1998; Bonaccorso *et al.* 2003). In the case of the application of the PCA to a normalized variable this means that the extracted PCs can be approximately considered representations of the same variable measured in the same units, which was assumed in this study.

As stated by Santos *et al.* (2010), compared to other regionalization methods there are some advantages in the use of PCA: (i) PCAs are not affected by the lack of independence in the original variables; (ii) normality is recommended but not an obligatory condition (Kalayci & Kahya 2006); and (iii) only an excessive number of zeros in the observations could cause problems, which in the applications envisaged is not a concern (Hair *et al.* 2005).

In the present study and since SPI is a normalized variable, accordingly with its calculation procedure, there was no need to previously transform the data.

Considering k variables in a given time period i, $X_{i,1}$, $X_{i,2}$, …, $X_{i,k}$, k PCs are produced for the same time period, $Y_{i,1}$, $Y_{i,2}$, …, $Y_{i,k}$, using linear combinations of the first ones, according to

$$\begin{cases} Y_{i,1} = a_{11}X_{i,1} + a_{12}X_{i,2} + \ldots + a_{1k}X_{i,k} \\ Y_{i,2} = a_{21}X_{i,1} + a_{22}X_{i,2} + \ldots + a_{2k}X_{i,k} \\ \quad \ldots \\ Y_{i,k} = a_{k1}X_{i,1} + a_{k2}X_{i,2} + \ldots + a_{kk}X_{i,k} \end{cases} \quad (2)$$

In the applications developed, the variables $X_{i,k}$ refer to the SPI series, k is equal to the number of rain gauges considered in the analysis and i represents the length of SPI series in each rain gauge.

In the previous combinations, the Y values or component scores (PC scores) are orthogonal and uncorrelated variables, such that $Y_{i,1}$ explains most of the variance, $Y_{i,2}$ the reminiscent amount of variance, and so on. The coefficients of the linear combinations are called 'loadings' and represent the weights of the original variables in the PCs.

PC extraction could be based on variance/covariance or correlation matrix of data with $\{a_{11}, a_{12}, \ldots, a_{1k}\}$ being the first eigenvector and $\{a_{k1}, a_{k2}, \ldots, a_{kk}\}$ being the eigenvector of k order. Each eigenvector includes the coefficients of the k principal component. In the present study the Pearson correlation matrix was considered for PC extraction.

Finally the amount of variance explained by the first PC is called the first eigenvalue, δ_1, the second is δ_2, so that $\delta_1 \geq \delta_2 \geq \delta_3 \geq \delta_4 \ldots \delta_k$, since each eigenvalue represents the fraction of the total variance in the original data and explained by each component (Bordi & Sutera 2001) so that this proportion can be calculated as $\delta_k / \sum \delta_k$. The analysis of the results of PCs can be focused on the eigenvalues (scree plot), on the correlations between PCs and the original variables (factor loadings) or on the percentage of the variance explained.

To achieve more stable spatial patterns, a rotation of the principal components with the Varimax procedure was applied. This procedure provides a clearer division between components, preserves their orthogonality and produces more physically explainable patterns (Richman 1986; Vicente-Serrano *et al.* 2004). Kahya *et al.* (2008a) referred that the rotation simplifies the spatial structure by isolating regions with similar temporal variations, being the Varimax procedure the most common orthogonal method to improve the creation of regions of maximum correlation between the variables and the components. The patterns defined in this way are referred to as rotated principal components (RPCs).

Changes in yearly drought occurrence

The analysis of changes in the temporal occurrences of droughts attempts to answer the question: regardless of the severity of the drought, i.e. the rainfall deficit, has the occurrence of droughts increased or decreased over time? Hence the analyzed variable is not related to the SPI themselves but to the temporal distribution of the occurrence of droughts.

To address the previous question, a KORE (Mudelsee *et al.* 2003; Silva *et al.* 2012) may be applied to a historical series of drought occurrences with the aim of estimating how the mean number of drought months in a year λ changes over time, that is to characterize $\lambda(t)$. This technique may be formulated as

$$\hat{\lambda}(t) = h^{-1} \sum_{i=1}^{m} K\left(\frac{t - T_i}{h}\right) \quad (3)$$

where K is the kernel function and h is the bandwidth. The following Gaussian kernel was used (Mudelsee *et al.* 2004, 2006)

$$K(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \quad (4)$$

The units of $\hat{\lambda}(t)$ are mo^{-1}, i.e. the average number of occurrences per month at any given time t. Nevertheless, in order to facilitate the interpretation of the results, $\hat{\lambda}(t)$ was multiplied by 12 such that, for any given instant t, $\hat{\lambda}(t)$ indicates the average number of occurrences per year (year^{-1}).

To reduce the boundary bias near the extremes of the time interval, pseudodata were generated outside of the observation interval, before estimating $\hat{\lambda}(t)$. For that purpose a straightforward method of reflection was used to generate pseudodata, covering the amplitude of three times the bandwidth h before and after the limits of the time interval (Mudelsee *et al.* 2004).

Other authors, such as Girardin & Mudelsee (2008), have also used the Kernel estimator for studying the occurrence rate of extreme events, namely the fire years in Canadian boreal forests. The results obtained from their approach, and using the same Kernel approach used herein, suggested that by the horizon 2061–2100, the median number of large forest fires per year could increase by 39% (ECHAM4 B2 scenario run) to 61% (A2 scenario run) when compared to the 1901–1940 and 1781–1820 reference periods used in the study. The same results if considering the full 1999–2100 horizon.

Regarding the use of drought indices, Christie *et al.* (2011) found for the Andes Cordillera that severe and extreme drought events reveal unprecedented increments on the respective occurrence rates during the last century when compared to the previous six centuries. For that purpose the previous authors considered the Palmer drought severity index to account with regional moisture and the same advanced statistical tool as in the present study, the Kernel estimation method, to analyze the occurrence rates of droughts.

To account for the uncertainty of the KORE estimates, a pointwise 90% confidence band was constructed around $\lambda(t)$, by means of a bootstrap simulations (Cowling *et al.* 1996; Mudelsee 2011), according to the methodology described in Silva *et al.* (2012). The KORE method coupled with bootstrap confidence band construction was first introduced into the analysis of climate extremes by Mudelsee *et al.* (2003), with a detailed description given by Mudelsee *et al.* (2004).

STUDY REGION AND RAINFALL DATA

In general terms, in South America the climate is predominantly wet and hot. However, the large size of the continent makes the climate vary, each region having its own characteristic weather conditions. Among the factors that influence climate are the orographic features, ocean currents and winds (Raia & Cavalcanti 2008).

The study region ranges from southern Brazil (the states of Paraná, Santa Catarina and Rio Grande do Sul) through Paraguay, to the contiguous northeastern states of Argentina (Formosa, Chaco, Corrientes, Misiones, Santiago del Estero, Santa Fé and Entre Ríos). It comprises an area of approximately 17.5 M km^2 and is characterized by pronounced rainfall variability (both in time and space). The different patterns of weather, climate and hence of rainfall variability over this region are among others factors in close connection with the long meridional orographic symbol of the continent, the Andes cordillera. This South American mountain system is also the world's longest, with a range that covers about 8,850 km and is situated on the far western edge of the continent, stretching from the southern tip to the northernmost coast of South America.

The El Niño Southern Oscillation phenomenon which has created a known ocean–atmosphere system in the tropical Pacific has a direct, strong influence over most of tropical and subtropical South America. Similarly, SST anomalies over the Atlantic Ocean also have a profound

impact on the climate and weather along the eastern coast of the continent (Garreaud & Aceituno 2001).

The data utilized in the analysis consisted of 51 series of monthly rainfall, from January 1961 to December 2011 (51 years), from rain gauges distributed on the study region. The monthly rainfall samples had a very few gaps that were filled by applying linear regression based on the records at nearby rain gauges, which is a common reconstruction technique of hydrologic time-series (Vicente-Serrano 2006). The names and geographical coordinates of the rain gauges, as well as the mean annual rainfall (MAR), are presented in Table 2. Figure 2 shows the study region, along with the location of the rain gauges.

The rainfall data from Paraguay and Argentina were obtained from the National Meteorological Services of both countries. In the case of Paraguay the institution is

Table 2 | Names, geographical coordinates and mean annual rainfall in the rain gauges

Rain gauge code	Country	Name	Longitude (°)	Latitude (°)	MAR (mm)
A1	Argentina	Santiado Del Estero	−64.30	−27.77	618
A2		Ceres	−61.95	−29.88	936
A3		Reconquista	−59.70	−29.18	1,239
A4		Sauce Viejo	−60.82	−31.70	999
A5		Parana	−60.48	−31.78	1,099
A6		Montes Caseros	−57.65	−30.27	1,456
A7		Concordia	−58.02	−31.30	1,355
A8		Rosario	−60.78	−32.92	1,022
A9		Gualeguaychú	−58.62	−33.00	1,103
A10		Las Lomitas	−60.58	−24.70	904
A11		Pcia.Roque Saenz Peña	−60.45	−26.82	1,078
A12		Resistencia	−59.05	−27.45	1,370
A13		Formosa	−58.23	−26.20	1,411
A14		Corrientes	−58.77	−27.45	1,428
A15		Posadas	−55.97	−27.37	1,754
A16		Paso de los Libres	−57.15	−29.68	1,527
B1	Brazil	Bagé	−54.10	−31.33	1,499
B2		Born jesus	−50.43	−28.67	1,728
B3		Campo Mourão	−52.37	−24.05	1,630
B4		Campos Novos	−51.20	−27.38	1,998
B5		Castro	−50.00	−24.78	1,571
B6		Caxias do Sul	−51.20	−29.17	1,798
B7		Chapeco	−52.62	−27.12	2,072
B8		Cruz Alta	−53.60	−28.63	1,813
B9		Curitiba	−49.27	−25.43	1,544
B10		Encruzilhada do Sul	−52.52	−30.53	1,600
B11		Florianópolis	−48.57	−27.58	1,671
B12		Idaial	−49.22	−26.90	1,738
B13		Iraí	−53.23	−27.18	1,898

(*continued*)

Table 2 | continued

Rain gauge code	Country	Name	Longitude (°)	Latitude (°)	MAR (mm)
B14		Irati	−50.63	−25.47	1,609
B15		Ivaí	−50.85	−25.00	1,663
B16		Lages	−50.33	−27.82	1,645
B17		Londrina	−51.13	−23.32	1,601
B18		Maringá	−51.92	−23.40	1,579
B19		Paranaguá	−48.52	−25.53	2,184
B20		Passo Fundo	−52.40	−28.22	1,849
B21		Porto Alegre	−51.17	−30.05	1,401
B22		Rio Grande	−52.10	−32.03	1,258
B23		Sta. Maria	−53.70	−29.70	1,719
B24		Sta. Vitória do Palmar	−53.35	−33.52	1,233
B25		S. Joaquim	−49.93	−28.30	1,793
B26		S. Luiz Gonzaga	−55.02	−28.40	1,911
B27		Torres	−49.72	−29.35	1,475
B28		Uruguaina	−57.08	−29.75	1,512
P1	Paraguay	Mcal. Estigarríbia	−60.61	−22.03	770
P2		Pto. Casado	−58.09	−22.21	1,245
P3		Concepción	−57.45	−23.35	1,356
P4		Asunción	−57.66	−25.18	1,394
P5		Villarrica	−56.46	−25.65	1,667
P6		Pilar	−58.31	−26.80	1,398
P7		Encarnación	−56.00	−27.20	1,765

the Direction of Meteorology and Hydrology of the National Directorate of Civil Aeronautical (DINAC – Spanish acronym; www.meteorologia.gov.py) while in Argentina it is the National Weather Service (www.smn.gov.ar) dependent of Defense Ministry of Argentina. In Brazil the data were provided by the National Institute of Meteorology (www.inmet.gov.br) under the Ministry of Agriculture, Livestock and Supply of Brazil.

RESULTS

Drought index and drought spatial patterns

Based on the monthly rainfall records at the rain gauges of Table 2 and Figure 2, the SPI series for the time-scales of 3 (SPI3) and 6 (SPI6) months were computed.

The characterization of the time evolution of the SPI at different time-scales in a given rain gauge is not a trivial task. One of the most suggestive and synthetic ways of performing such characterization is exemplified in Figure 3, the concept taken from http://joewheatley.net/visualizing-drought/. The figure illustrates the time evolution of the SPI at the time-scales from 1 to 12 months at one of the rain gauges analyzed in this study, and further characterized – the rain gauge A1 – Las Lomitas, located in northern Argentina. We note that for producing Figure 3, SPI1–12 were calculated, notwithstanding that the paper generally focuses on SPI3 and SPI6 on the other analyses.

The previous figure easily stresses the pronounced recurrence of dry periods in the 1960s and 1970s (yellow/brown colors especially marked in the higher time-scales of SPI), followed by a wetter period (green/blue colors) in the 1980s and early 1990s. Figure 2 also shows that for small

time-scales (e.g. 3 months) each new month added in the SPI calculation has a large impact on the period sum of precipitation, so it is relatively easy and more frequent to have the SPI responding quickly and moving from dry to wet values, in consequence we have more events. For bigger time-scales (e.g. 12 months) each new month added into the calculation has less impact on the total and the SPI responds more slowly, which leads to fewer droughts but with longer duration (McKee *et al.* 1993).

To identify the spatial patterns of drought at different time-scales in the study area, PCA was applied. According to Equation (2) the variables $X_{i,k}$ refer to the SPI series, k is the number of rain gauges (51) and i represents the length of SPI series in each rain gauge. For SPI at 3 and 6 months (SPI3 and SPI6), i varies from 1 to $51 \times 12^{-2} = 610$ and from 1 to $51 \times 12^{-5} = 607$, respectively.

For both SPI time-scales, the number of principal components retained for Varimax rotation was selected based on the interpretation of the scree plot (Bryant & Yarnold 1995), on the mapping of the factor loadings (raw data) and on the amount of variance explained in the original data.

On the basis of 51 years of monthly rainfall in the 51 rain gauges schematically located in Figure 2, 51 series of SPI were obtained for each time-scale. As previously mentioned, the SPI series have 610 and 607 accumulated rainfalls for the time-scales of 3 (SPI3) and 6 (SPI6) continuous months, respectively.

According to the examination of Figure 4, the scree plot shows that the line stops descending precipitously and levels out approximately on the fourth PC, which gives an indication to retain between three or four principal components. Taking into account the variance explained by each component, Figure 5 shows the first 10 PCs retained, being clear that the first seven components explain about 68% of the total variance in the original SPI series, for the SPI3, and about 69% for the SPI6. It could be noted that from PC eight to 10 these components explain, when compared to the first ones, only a small amount of variance in the original SPI data (2% each of the total variance).

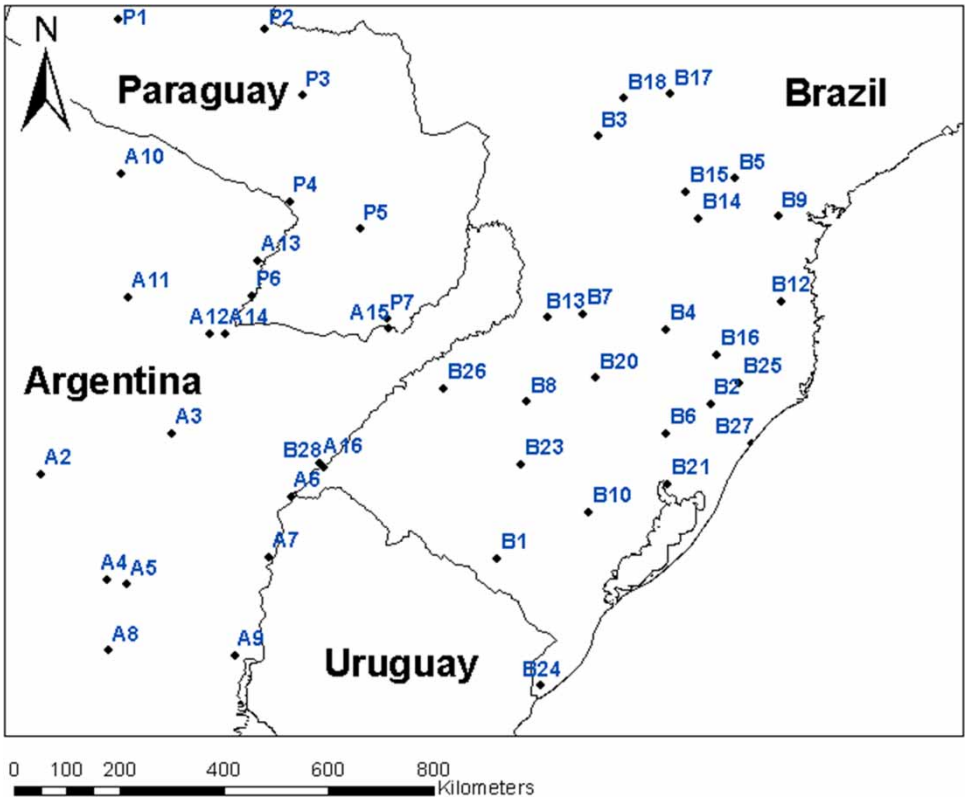

Figure 2 | Map of the study area and location of the 51 rain gauges.

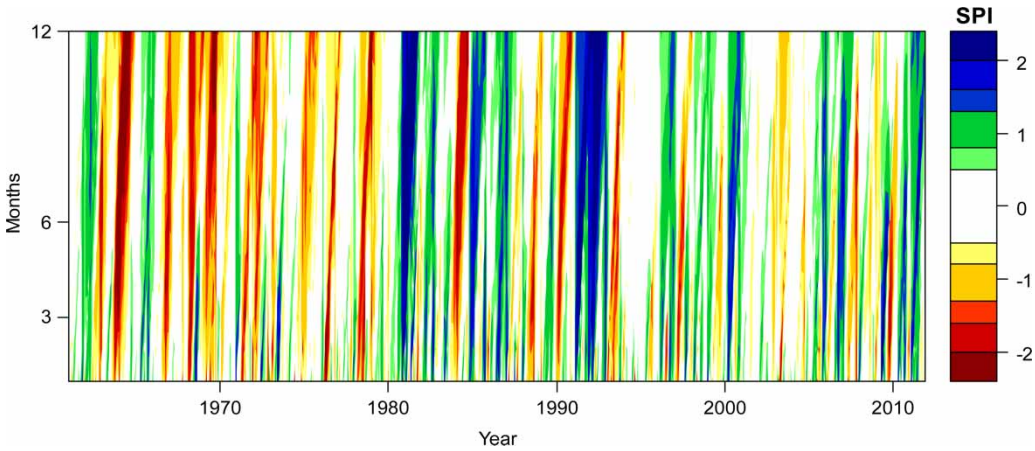

Figure 3 | Representation of the time evolution of the SPI at several time-scales. Example based on the rain gauge A1 – Las Lomitas.

Seven leading components were also suggested from the analysis of the mapping of the factor loadings (correlations between the original data – SPI series at 51 rain gauges – and the PC series scores), since they fully cover the study area and do not overlap (similar patterns from those of Figures 6 and 7).

Based on the results, and for the purpose of regionalization, which means choosing the number of components to rotate and include in the PCA, we agreed to retain the seven main patterns or RPCs that were considered in the study, F1–F7.

The spatial extent of the first seven retained RPCs (components F1–F7) that covers the entire study region was characterized by mapping the raw values of the factor loadings, since they are an important indicator to identify the region that can be correlated to a specific RPC.

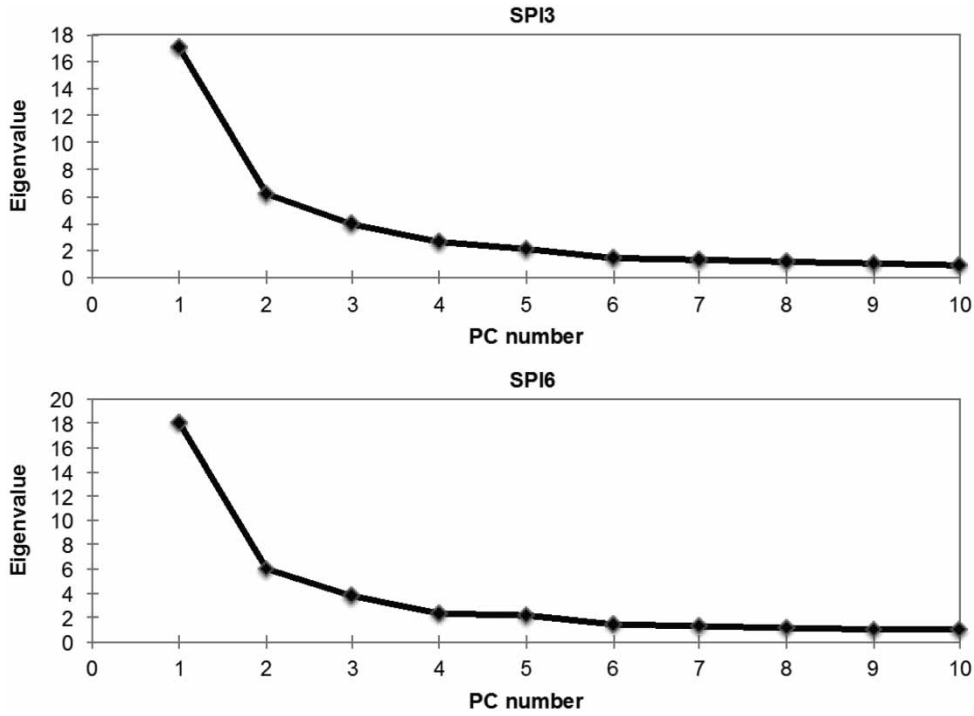

Figure 4 | First 10 eigenvalues resulting from the PCA applied to the SPI computed on both time-scales (SPI3 and SPI6).

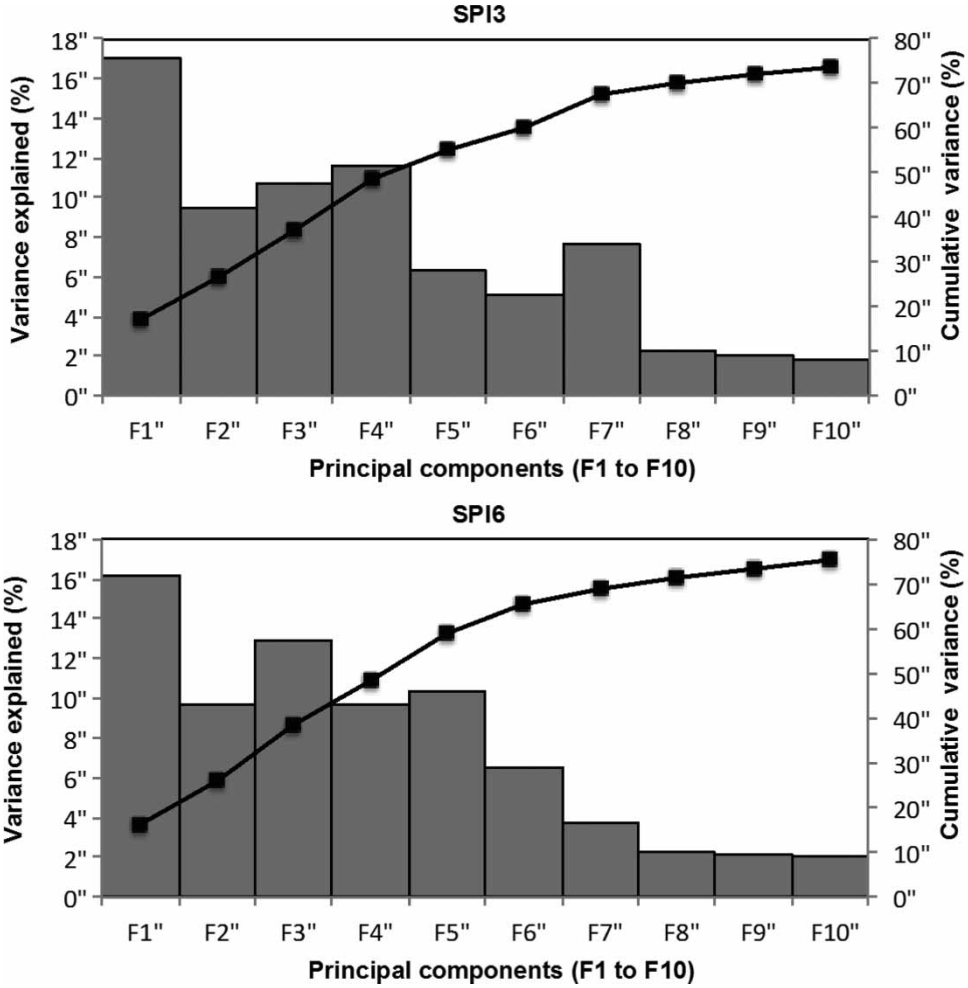

Figure 5 | Percentage of variance explained by first 10 components of PCA, F1–F10.

For that purpose, the Kriging spatial interpolation method (Oliver & Webster 1990) available on Arcgis version 10.1 (http://www.esri.com/software/arcgis/arcgis-for-desktop) was utilized. The results achieved are represented in Figures 6 and 7 along with the results from the KORE frequency estimator applied to each one of the RPC time-series (F1–F7).

Figures 6 and 7 show that between the first seven components, F1–F7, the regions with significant statistical correlation between the main RPCs and the original SPI time-series field do not overlap, being clearly spatially disjunctive. More precisely in each of the SPI time-scales the first five components (F1–F5) showed positive correlations that were always higher than 0.7. This means that only a specified number of rain gauges are the most important for each RPC extraction and therefore they also constitute the main representatives of regional drought behavior.

According to Yevjevich (1972), for samples such as those analyzed with 610 and 607 observations (months), linear correlations of 0.07 are significant at a confidence level of 5%.

For the SPI3 time-scale, the first component (F1) highlights an area located in the southern part of Brazil and it explains nearly 17% of the total variance while for the SPI6 the same first component (F1) explains around 16% (Figure 5). The second component (F2) for the SPI3 time-scale explains an area in the southern part of Argentina, and the third (F3) in the northern part of the country. The fourth and the fifth components (F4 and F5) explain areas in Brazil, in the north and eastern coastal part, respectively.

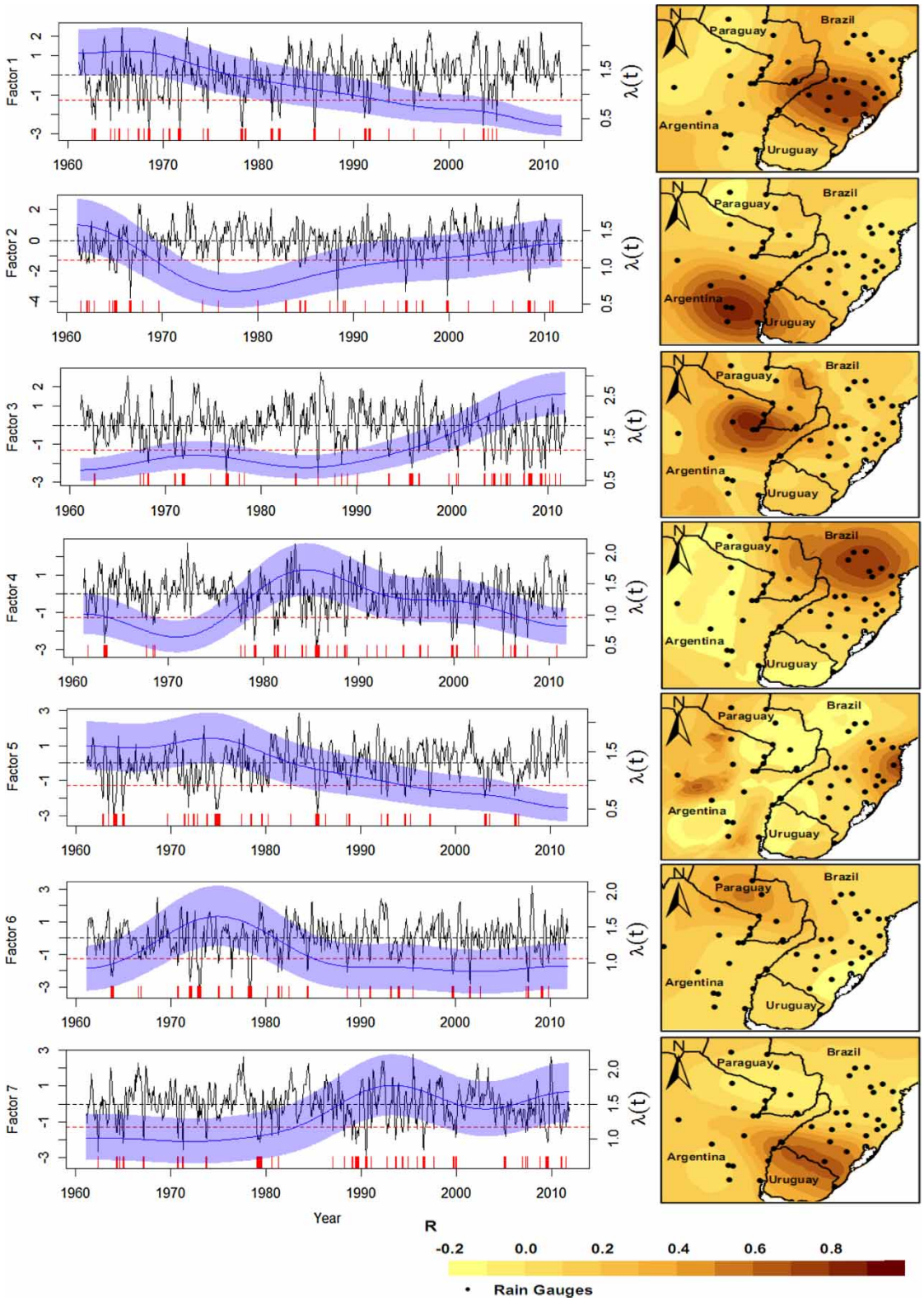

Figure 6 | Time-dependent occurrence rates of severe drought months of the seven RPCs of SPI3 (left); spatial correlation maps between each RPC and the at-site SPI series (right). The vertical ticks on x-axis indicate the points in time when drought events occurred.

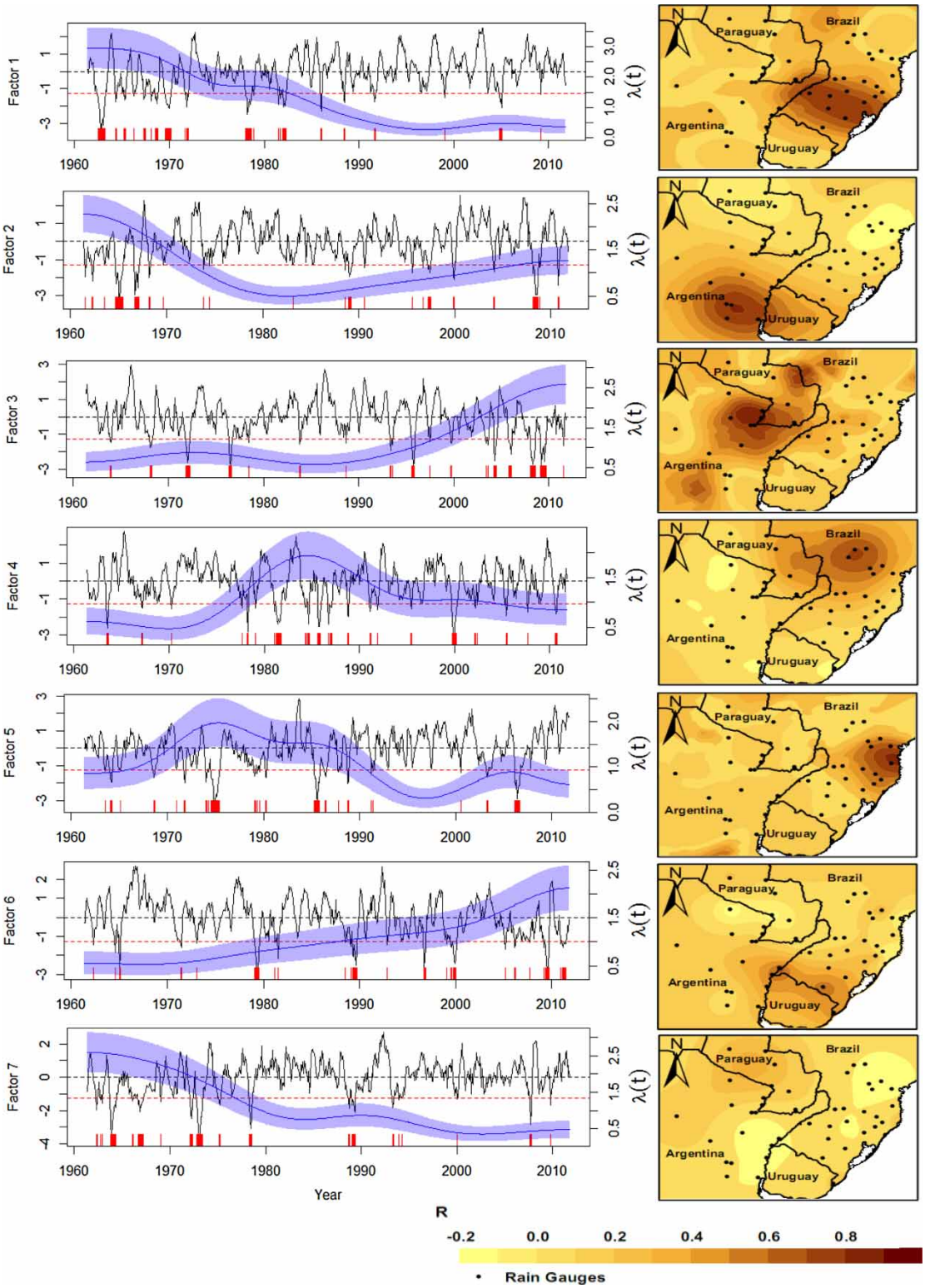

Figure 7 | Time-dependent occurrence rates of severe drought months of the seven RPCs of SPI6 (left); spatial correlation maps between each RPC and the at-site SPI series (right). The vertical ticks on x-axis indicate the points in time when drought events occurred.

Lastly the sixth and seventh components (F6 and F7) explain areas in the center of Paraguay and in the Brazilian southern border with Uruguay. For the SPI6 time-scale the results are equivalent with the exception that the sixth most important component (F6) identifies the same region as the seventh component (F7) in the SPI3 time-scale, and vice versa.

All of the rotated components relate mostly positively with the original SPI series, the negative correlations being quite localized on the maps and with little statistical significance (R with a minimum of -0.2). Retaining seven PCs means that the variation measured by the SPI among the drought/wet conditions across the entire study region at any given time can be explained adequately by seven components, rather than 51 rain gauges (dimensionality reduction of the SPI field). The results show that a regionalization could be achieved, the identified regions being similar for both SPI3 and SPI6.

Changes in drought occurrence rate

To analyze the changes in the yearly drought occurrence rate the KORE estimator was applied to the seven RPCs, F1–F7, previously obtained for each time-scale considered (3 and 6 months). The analysis focused on the occurrence of severe droughts, that is the occurrence of SPI values lower than -1.28, represented by the vertical ticks on the x-axes, included in Figures 6 and 7 (point process). The bandwidth h that appears in Equation (3) was obtained using Silverman's 'rule of thumb', Silverman (1986: 48). Its values range from 1.387 to 7.513 years.

Figures 6 and 7 show the KORE estimates and the associated bootstrap confidence band to the SPI3 and SPI6, respectively, of the seven RPCs together with maps of correlation between the at-site SPI series with each RPC.

Comparison of the results of Figures 6 and 7 for the same RPC shows that the temporal behavior of the severe drought occurrence rate is moderately spatially coherent across the region, i.e., is relatively independent of the SPI time-scale. For the northeastern Argentina region (area highlighted by RPC3), the drought occurrence rate shows some increasing tendency from the mid-1980s, while for southern Brazil (RPC1) there seems to be a decreasing trend. Though much less evident, a decrease in the drought frequency seems to occur in Paraguay (RPC6 and RPC7 for the time-scale of 3 and 6 months, respectively) and an increase in the region along the border between Brazil and Paraguay (RPC7 and RPC6 for those same time-scales). The remaining sub regions highlighted by each RCP do not show an observable trend when both SPI time-scales are considered.

For a region in central Argentina located near the area identified by RPC2 (Figures 6 and 7), Capriolo & Scarpati (2012), using a different approach based on soil water balance coupled with the Penman–Monteith evapotranspiration method, found increasing drought trends in the periods 1991–2010 and 2001–2010, which were statistically significant at level $\alpha = 0.1$. Drought events were identified and quantified in respect to soil water deficits (agricultural drought) calculated on a monthly basis. These results could be comparable with our drought occurrence approach since we use the agricultural drought definition related to SPI (3- and 6-month time-scales). In fact, they are in agreement with the results here achieved of the time-dependent occurrence rates (KORE) of severe drought months for SPI3 and SPI6 in the RPC 2, where we see a notable increase of the drought occurrence rate from the 1980s to the end of 2011. The same authors did not find statistically significant trends for drought during the total studied period (1971–2010).

Rainfall thresholds for drought recognition

Despite the widespread use and the advantages of SPI compared to other drought indices, the interpretation of the values associated with SPI and drought monitoring based on those values are not easy to accomplish, especially as they involve standardized values that are difficult to relate the rainfall with which mathematical manipulation they result.

Therefore, an additional calculation was developed that gives the SPI values that represent drought thresholds back to the rainfall field, thus facilitating an adequate interpretation of the meaning of such index and quite easily and reliably identifying the drought episodes (Portela *et al.* 2012). As a result, monitoring can be operationalized as can the subsequent actions that need to be undertaken.

For that purpose and for all the 51 rain gauges of Table 2, the cumulative rainfall in 3 and 6 consecutive months were estimated for an arbitrary value of SPI of

−1.65 (extreme drought, that according to Agnew (2000) is associated to a 5% historical occurrence, Table 1). The previous estimation required the inversion of the SPI calculation procedure through the use of a set of widely tested computational subroutines (Hosking 1996).

Figures 8 and 9 exemplify the results obtained for the previous drought threshold and for SPI3 and SPI6, respectively. Each one of the maps that appears in these figures shows the spatial distribution of the rainfall in the groups of months identified in the body of the map. Whenever the rainfall registered in a given location and period falls below the value shown by the map for such a location and period, then the location is suffering an extreme drought. Therefore, each map represents the surface of the

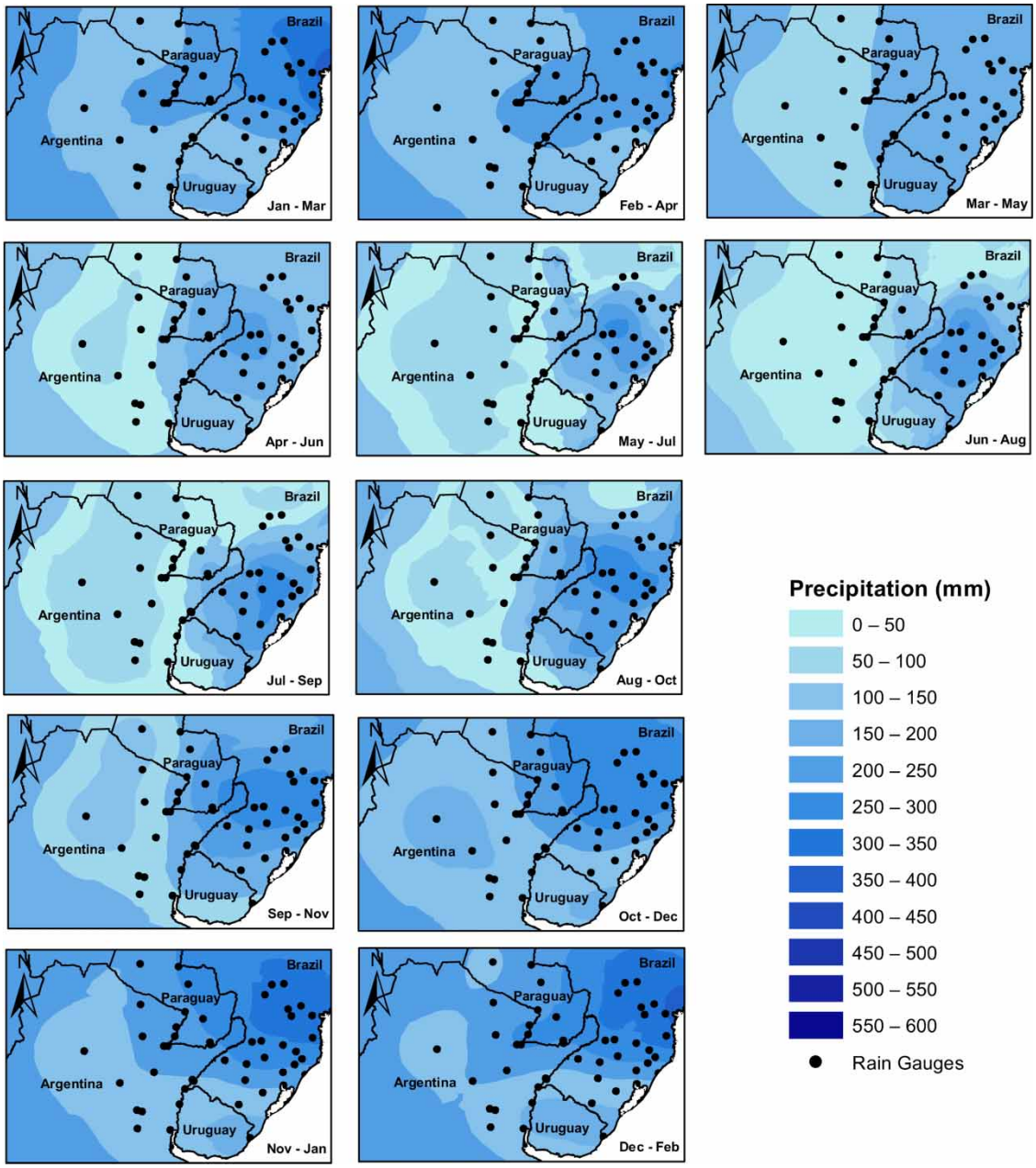

Figure 8 | Inversion of SPI3 = − 1.65. Rainfall in 3 consecutive months corresponding to the extreme drought threshold.

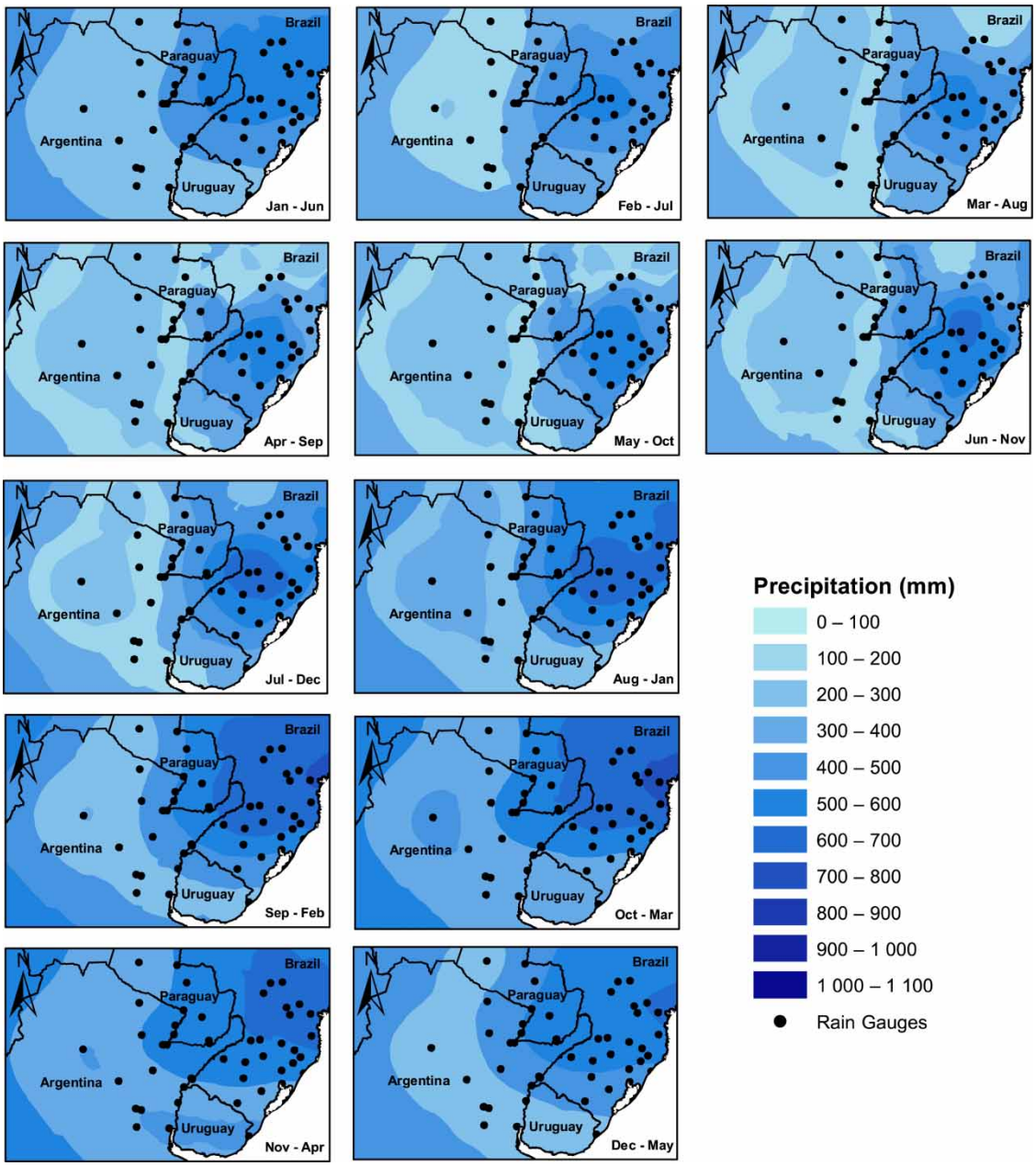

Figure 9 | Inversion of SPI6 = – 1.65. Rainfall in 6 consecutive months corresponding to the extreme drought threshold.

minimum rainfall in a given interval of time below which an extreme drought episode is recognized – surfaces of cumulative rainfall thresholds for extreme drought.

The rainfall surfaces were achieved by applying the Kriging spatial interpolation geostatistical method (Oliver & Webster 1990) to the rainfall thresholds obtained by the inversion of the SPI at different time-scales in the 51 rain gauges that supported the study (Table 2). The schematic localization of the gauges was also included in Figures 8 and 9.

The maps included in each figure provide the cumulative rainfall in k months (with k equal to 3 or 6 in correspondence with the SPI time-scale, i.e. SPI3 and SPI6), beginning in the successive months of the civil year.

In this way, the result for each time-scale (3 or 6 months) is always 12 maps. The period of consecutive months to which the rainfall thresholds refer is identified in each map.

These figures show very smooth rainfall patterns, especially for the time-scale of 3 months (Figure 8). Such a circumstance is particularly evident for the cumulative rainfall from March to May: where both interior and coastal regions have shown different responses to extreme drought for the same period of consecutive months to which the rainfall thresholds refer. In fact, those regions simply show different climates, which is one of the benefits of normalization when using the SPI. Since coastal regions have more annual rainfall then their fifth percentile is higher than the drier, inland areas. Drought refers to an anomaly from a reference value (e.g. mean annual rainfall), so for the same rainfall amount the drier inland regions are more accustomed and have less drought effects compared to the coastal areas.

DISCUSSION AND CONCLUSIONS

An exploratory approach aiming at characterizing the spatial pattern and the occurrence rates of droughts in Paraguay, southern Brazil and northeastern Argentina was carried out. To identify the drought events, the SPI was applied to the monthly rainfall series at 51 rain gauges at 3- and 6-month time-scales (3 and 6 consecutive months for SPI3 and SPI6, respectively).

The regionalization resulting from the PCA allowed identifying seven contiguous and non-overlapping regions with the same drought patterns. The regionalization was similar for both SPI3 and SPI6, which points toward a stable and consistent spatial pattern across other SPI time-scales.

Regarding the analysis of the drought occurrences rates, the achieved results for both time-scales, despite proving the suitability of the approach based on the KORE method coupled with bootstrap confidence band construction, only allowed us to clearly identify two regions where trends in the drought frequency seems to occur, with opposite signals: northeastern Argentina and Brazil, the first with more droughts and the second with fewer droughts up to the present day.

When analyzing the results one should keep in mind the reduced number of rain gauges that were considered coupled with the relatively small length of the rainfall samples. Accordingly, the next step of the study should expand the procedure by applying it to a denser network of rain gauges with an extended period of records, which will also allow a more detailed characterization of the droughts in the study area.

Despite its simplicity, the surfaces of cumulative rainfall thresholds for drought recognition can be a very helpful tool for drought management as they allow, in a timely and simple way, to recognize drought events and to follow their evolution. As a result, monitoring can be operationalized as can the subsequent actions that need to be undertaken. Therefore, more surfaces for other SPI time-scales and drought categories should be obtained, especially when more rainfall data are available.

Water resources managers should also take in account these surfaces in order to alleviate possible future water shortages that could become more frequent and can cause significant economic losses in the agricultural sector of southern South America.

REFERENCES

Abdi, H. & Williams, L. J. 2010 Principal component analysis. *WIREs Comput. Stat.* **2**, 433–459.

Agnew, C. T. 2000 Using the SPI to identify drought. *Drought Network News* **12**, 6–12.

Barrucand, M. G., Vargas, W. M. & Rusticucci, M. M. 2007 Dry conditions over Argentina and the related monthly circulation patterns. *Meteorol. Atmos. Phys.* **98** (1–2), 99–114, doi:10.1007/s00703-006-0232-5.

Bonaccorso, B., Bordi, I., Cancelliere, A., Rossi, G. & Sutera, A. 2003 Spatial variability of drought: an analysis of the SPI in Sicily. *Water Resour. Manage.* **17**, 273–296.

Bordi, I. & Sutera, A. 2001 Fifty years of precipitation: some spatially remote teleconnections. *Water Resour. Manage.* **15**, 247–280.

Bryant, F. B. & Yarnold, P. R. 1995 Principal-components analysis and confirmatory factor analysis. In: *Reading and Understanding Multivariate Statistics* (L. G. Grimm & P. R. Yarnold, eds). American Psychological Association, Washington, DC.

Capriolo, A. D. & Scarpati, O. E. 2012 Extreme hydrologic events in north area of Buenos Aires Province (Argentina). *ISRN Meteorol.* **2012**, 415081.

Changnon, S. A. & Easterling, W. E. 1989 Measuring drought impacts: the Illinois case. *Water Resour. Bull.* **25**, 27–42.

Christie, D. A., Boninsegna, J. A., Cleaveland, M. K., Lara, A., Le Quesne, C., Morales, M. S. & Mudelsee, M. 2011 Aridity changes in the temperate-Mediterranean transition of the Andes since AD 1346 reconstructed from tree-rings. *Clim. Dyn.* **36**, 1505–1521.

Cowling, A., Hall, P. & Phillips, M. 1996 Bootstrap confidence regions for the intensity of a Poisson point process. *J. Am. Stat. Assoc.* **91**, 1516–1524.

Edwards, D. C. & McKee, T. B. 1997 *Characteristics of 20th Century Drought in the United States at Multiple Time Scales*. Climatology Report 97–2, Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado.

Ehrendorfer, M. 1987 A regionalization of Austria's precipitation climate using principal component analysis. *J. Climatol.* **7**, 71–89.

Fraisse, C. W., Cabrera, V. E., Breuer, N. E., Baez, J., Quispe, J. & Matos, E. 2008 El Niño-southern oscillation influences on soybean yields in eastern Paraguay. *Int. J. Climatol.* **28**, 1399–1407.

Garreaud, R. D. & Aceituno, P. 2001 Atmospheric circulation over South America: mean features and variability. In: *The Physical Geography of South America* (T. Veblen, K. Young & A. Orme, eds). Oxford University Press, New York, Chap. 2.

Girardin, M. P. & Mudelsee, M. 2008 Past and future changes in Canadian boreal wildfire activity. *Ecol. Appl.* **18**, 391–406.

Grimm, A. M. 2004 How do La Niña events disturb the summer monsoon system in Brazil? *Clim. Dyn.* **22**, 123–138.

Grimm, A. M., Barros, V. R. & Doyle, M. E. 2000 Climate variability in southern South America associated with El Niño and La Niña events. *J. Clim.* **13**, 35–58.

Guttman, N. B. 1998 Comparing the palmer drought index and the standardized precipitation index. *J. Am. Water Resour. Assoc.* **34**, 113–121.

Guttman, N. B. 1999 Accepting the standardized precipitation index: a calculation algorithm. *J. Am. Water Resour. Assoc.* **35**, 311–322.

Hair, J. F., Anderson, R. E., Tatham, R. L., Babin, B. & Black, B. (eds) 2005 *Multivariate Data Analysis*, 6th edn. Prentice-Hall, London. 928 pp.

Hayes, M., Wilhite, D. A., Svoboda, M. & Vanyarkho, O. 1999 Monitoring the 1996 drought using the standardized precipitation index. *Bull. Am. Meteorol. Soc.* **80**, 429–438.

Hosking, J. R. M. 1996 Fortran code written for inclusion in IBM research report RC20525. In: *Fortran Routines for Use with Method of L-moments*. IBM Research Division, T. J. Watson Research Center, New York, USA.

Inter-American Development Bank 2014 Country Program Evaluation Paraguay 2009–2013 Office of Evaluation and Oversight, Washington, USA.

Jolliffe, I. T. (ed.) 2002 *Principal Component Analysis*, 2nd edn. Springer, New York, 502 pp.

Kahya, E., Demirel, M. C. & Beg, O. A. 2008a Hydrologic homogeneous regions using monthly streamflow in Turkey. *Earth Sci. Res. J.* **12**, 181–193.

Kahya, E., Kalayc, S. & Piechota, T. C. 2008b Streamflow regionalization: case study of Turkey. *J. Hydrol. Eng.* **13**, 205–214.

Kalayci, S. & Kahya, E. 2006 Assessment of streamflow variability modes in Turkey: 1964–1994. *J. Hydrol.* **324**, 163–177.

Lee, S. M., Byun, H. R. & Tanaka, H. L. 2012 Spatiotemporal characteristics of drought occurrences over Japan. *J. Appl. Meteorol. Climatol.* **51**, 1087–1098.

Lins, H. F. 1985 Interannual streamflow variability in the United States based on principal components. *Water Resour. Res.* **21**, 691–701.

Lloyd-Hughes, B. 2002 *The Long-Range Predictability of European Drought*. PhD Thesis (submitted), University of London, UK.

Lloyd-Hughes, B. & Saunders, M. A. 2002 European drought climatology and prediction using the standardized precipitation index (SPI). In: *8.11 13th Conference on Applied Meteorology*, 13–16 May 2002, Amer. Meteor. Soc., Portland, OR (USA).

Masuda, T. & Goldsmith, P. D. 2009 World soybean production: area harvested, yield, and long-term projections. International food and agribusiness management association (IAMA). *Int. Food Agribusiness Manage. Rev.* **12** (4), 143–162.

Mattos, J. G., Santos, A. F., Gonçalves, L. G. & Herdies, D. L. 2012 Estimating drought episodes over South Brazil. In: *Proceedings of 8th Alexander von Humboldt International Conference*, 12–16 November 2012, Cusco, Peru.

McKee, T. B., Doesken, N. J. & Kleist, J. 1993 The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference on Applied Climatology, American Meteorology Society, 17–22 January, Anaheim, CA*. American Meterological Society, Boston, MA, pp. 179–184.

Mudelsee, M. 2011 The bootstrap in climate risk analysis. In: *Extremis: Disruptive Events Trends in Climate and Hydrology* (J. P. Kropp & H. J. Schellnhuber, eds). Springer, New York, pp. 45–58.

Mudelsee, M., Borngen, M., Tetzlaff, G. & Grunewald, U. 2003 No upward trends in the occurrence of extreme floods in central Europe. *Nature* **425**, 166–169.

Mudelsee, M., Borngen, M., Tetzlaff, G. & Grunewald, U. 2004 Extreme floods in central Europe over the past 500 years: role of cyclone pathway Zugstrasse Vb. *J. Geophys. Res.* **109**, D23101.

Mudelsee, M., Deutsch, M., Borngen, M. & Tetzlaff, G. 2006 Trends in flood risk of the River Werra (Germany) over the past 500 years. *Hydrol. Sci. J.* **51** (5), 818–833.

Ntale, H. K. & Gan, T. 2003 Drought indices and their application to East Africa. *Int. J. Climatol.* **23**, 1335–1357.

Oliver, M. A. & Webster, R. 1990 Kriging: a method of interpolation for geographical information system. *Int. J. Geogr. Inf. Syst.* **4**, 313–332.

Portela, M. M., Santos, J. F., Naghettini, M., Matos, J. P. & Silva, A. T. 2012 Superfícies de limiares de precipitação para

identificação de secas em Portugal continental: uma aplicação complementar do Índice de Precipitação Padronizada, SPI. *Recursos Hídricos* **33**, 5–23, Associação Portuguesa dos Recursos Hídricos, (APRH), Lisboa. http://www.aprh.pt/rh/v33n2.html.

Raia, A. & Cavalcanti, I. F. A. 2008 The life cycle of the South American monsoon system. *J. Clim.* **21**, 6227–6246.

Ravelo, A. C. 2000 Agroclimatic characterization of extreme droughts in the Pampas region of Argentina. *Revista de la Facultad de Agronomía (Universidad de Buenos Aires)* **20**, 187–192.

Rencher, A. C. (ed.) 1998 *Multivariate Statistical Inference and Applications*, John Wiley, New York.

Richman, M. B. 1986 Rotation of principal components. *J. Climatol.* **6**, 29–35.

Rossi, G. 2003 Requisites for a drought watch system. In: *Tools for Drought Mitigation in Mediterranean Regions* (G. Rossi, ed.). Kluwer Academic Publishing, Dordrecht, pp. 147–157.

Santos, J. F. & Portela, M. M. 2010 Caracterização de secas em bacias hidrográficas de Portugal Continental: aplicação do índice de precipitação padronizada, SPI, a séries de precipitação e de escoamento. In: 10° Congresso da Água, Associação Portuguesa dos Recursos Hídricos (APRH), Alvor, 18 pp.

Santos, J. F., Pulido-Calvo, I. & Portela, M. M. 2010 Spatial and temporal variability of droughts in Portugal. *Water Resour. Res.* **46**, W03503.

Santos, J. F., Portela, M. M. & Pulido-Calvo, I. 2011 Regional frequency analysis of droughts in Portugal. *Water Resour. Manage.* **25**, 3537–3558.

Santos, J. F., Portela, M. M., Naghettini, M., Matos, J. P. & Silva, A. T. 2013 Precipitation thresholds for drought recognition: a complementary use of the SPI. In: *7th International Conference on River Basin Management including all aspects of Hydrology, Ecology, Environmental Management, Flood Plains and Wetlands, RBM13*, Wessex Institute, New Forest, UK.

Sheffield, J. & Wood, E. F. 2008 Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Clim. Dyn.* **31**, 79–105.

Silva, A. T., Portela, M. M. & Naghettini, M. 2012 Nonstationarities in the occurrence rates of flood events in Portuguese watersheds. *Hydrol. Earth Syst. Sci.* **16**, 241–254.

Silverman, B. W. 1986 *Density Estimation for Statistics and Data Analysis, Monographs on Statistics and Applied Probability*, Vol. 26. Chapman & Hall/CRC, New York.

Sims, A. P., Nigoyi, D. S. & Raman, S. 2002 Adopting indices for estimating soil moisture: a North Carolina case study. *Geophys. Res. Lett.* **29**, 1183.

Singh, P. K., Kumar, V., Purohit, R. C., Kothari, M. & Dashora, P. K. 2009 Application of principal component analysis in grouping geomorphic parameters for hydrologic modelling. *Water Resour. Manage.* **23**, 325–339.

Szalai, S. & Szinell, C. 2000 Comparison of two drought indices for drought monitoring in Hungary – a case study. In: *Drought and Drought Mitigation in Europe* (J. V. Vogt & F. Somma, eds). Kluwer Academic Publishers, Dordrecht, pp. 161–166.

Tallaksen, L. M. & Van Lanen, H. A. 2004 *Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater*, Vol. 48. Elsevier Science, Amsterdam.

Tate, E. L. & Gustard, A. 2000 Drought definition: a hydrological perspective. In: *Drought and Drought Mitigation in Europe* (J. V. Vogt & F. Somma, eds). Kluwer Academic Publishers, Dordrecht, pp. 23–48.

Tipping, M. E. & Bishop, C. M. 1999 Probabilistic principal component analysis. *J. R. Stat. Soc., Ser. B* **61**, 611–622.

Vicente-Serrano, S. M. 2006 Spatial and temporal analysis of droughts in the Iberian Peninsula (1910–2000). *Hydrol. Sci. J.* **51**, 83–97.

Vicente-Serrano, S. M., González-Hidalgo, J. S., Luis, M. & Raventós, J. 2004 Drought patterns in the Mediterranean area: the Valencia region (eastern Spain). *Clim. Res.* **26**, 5–15.

Vogt, J. V. & Somma, F. 2000 *Drought and Drought Mitigation in Europe*. Kluwer Academic Publishers, Dordrecht, 325 p.

Westra, S., Brown, C., Lall, U. & Sharma, A. 2007 Modeling multivariable hydrological series: principal component analysis or independent component analysis? *Water Resour. Res.* **43**, W06429.

Wilhite, D. A. 1994 *Preparing for Drought: A Guidebook for Developing Countries*. DIANE Publishing, USA.

Wilhite, D. A. & Glantz, M. H. 1985 Understanding the drought phenomenon: the role of definitions. *Water Int.* **10**, 111–120.

Yevjevich, V. 1972 *Stochastic Processes in Hydrology*. Water Resources Publications, Fort Collins, CO.

First received 10 April 2014; accepted in revised form 18 October 2014. Available online 27 November 2014