Theoretical and Applied Climatology
https://doi.org/10.1007/s00704-019-02933-6

ORIGINAL PAPER

Summer seasonal predictability of warm days in Argentina: statistical model approach

Soledad Collazo [1,2] · Mariana Barrucand [1,2] · Matilde Rusticucci [1,2]

Received: 4 July 2018 / Accepted: 24 June 2019

Abstract

Predicting extreme temperature events can be very useful for different sectors that are strongly affected by their variability. The goal of this study is to analyze the influence of the main atmospheric, oceanic, and soil moisture forcing on the occurrence of summer warm days and to predict extreme temperatures in Argentina northern of 40°S by fitting a statistical model. In a preliminary analysis, we studied trends and periodicities. Significant positive trends, fundamentally in western Argentina, and two main periodicities of summer warm days were detected: 2–4 years and approximately 8 years. Lagged correlations allowed us to identify the key predictors: El Niño-Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO), and Standardized Precipitation Indices (SPI). We also noticed that the frequency of warm days in spring acts as a good predictor of summer warm days. Due to the collinearity among many predictors, principal component regression was used to simulate summer warm days. We obtained negative biases (i.e., the model tends to underestimate the frequency of summer warm days), but the observed and simulated values of summer warm days were significantly correlated, except in northwest Argentina. Finally, we analyzed the predictability of the summer warm days under ENSO neutral conditions, and we found new predictors: the geopotential height gradient in 850 hPa (between the Atlantic Anticyclone and the Chaco Low) and the Atlantic Multidecadal Oscillation (AMO), while the PDO and SPI lost some relevance.

1 Introduction

Extreme events have a higher impact than mean climate in sectors more closely linked, such as agriculture, human health, and water resources, among others. As highlighted in the Special Report on Managing the Risks of Extreme Events of the Intergovernmental Panel on Climate Change (IPCC 2012), extreme climate events and their changes are especially relevant to society and ecosystem due to their potentially dangerous impacts. The ability to predict these extremes would allow the undertaking of precautionary measures to avoid or reduce their impacts (Mueller and Seneviratne 2012). In particular, we focused in the summer because hot extremes have the highest impacts on human health and the environment during this time of year (Rusticucci et al. 2016).

In the summer season, the basic South America low-level atmospheric circulation features (Fig. 1) consist of the subtropical South Pacific anticyclone (SPA), the subtropical South Atlantic anticyclone (SAA), the "Chaco" continental orographic–dynamic–thermal low (CHL; Seluchi et al. 2003), the easterlies over tropical–equatorial Atlantic and Amazonia, and the mid-latitude westerlies. The easterlies are naturally deflected by the Andes orography towards the southeast and south into subtropical latitudes producing a north–south low-level jet (Agosta and Compagnucci 2008). During the Chaco low-level jet events, there are a large moisture flux and convergence at low and mid levels which are responsible for a significant fraction of the summer precipitation (Salio et al. 2002). The mean maximum temperature shows the highest values in Paraguay and northern Argentina, due to the incursion of air masses of tropical origin into midlatitudes, and it is frequently related to the Chaco and northwestern Argentine depression (Schwerdtfeger 1976).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00704-019-02933-6) contains supplementary material, which is available to authorized users.

✉ Soledad Collazo
scollazo@at.fcen.uba.ar

[1] Department of Atmospheric and Ocean Sciences, Faculty of Exact and Natural Sciences, University of Buenos Aires (DCAO-FCEN-UBA), Intendente Güiraldes 2160, Ciudad Universitaria Pab II, 1428 Buenos Aires, Argentina

[2] National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina

Published online: 13 July 2019

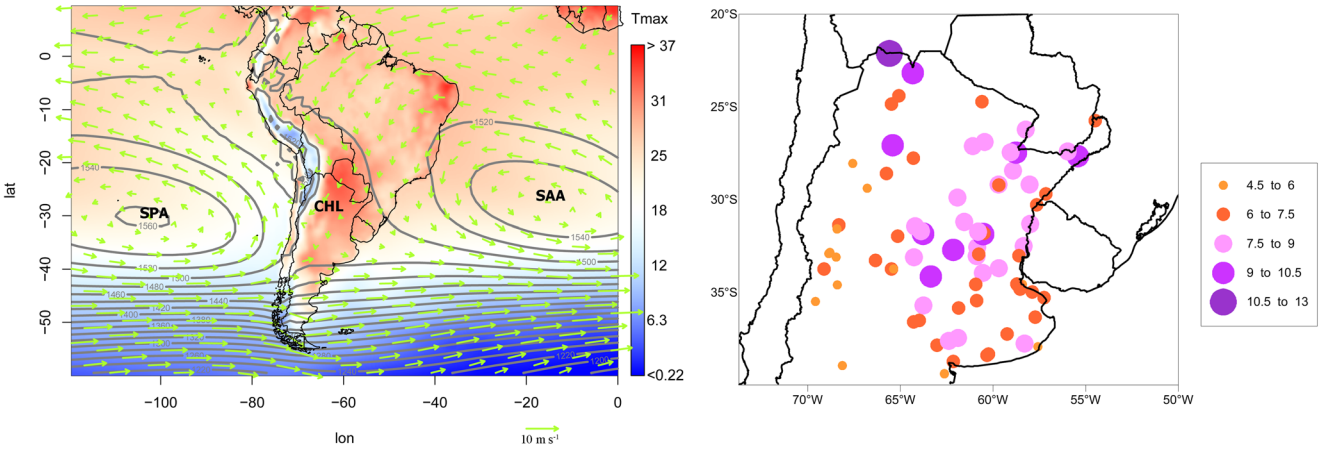

Fig. 1 Basic low-level atmospheric circulation features in the austral summer season (left): the summer mean maximum temperature is shaded, the contours are the seasonally averaged 850-hPa geopotential height, and the arrows are the 850-hPa vector wind, from the ERA-Interim. Climatology 1981–2010. SPA: subtropical South Pacific anticyclone. SAA: subtropical South Atlantic anticyclone. CHL: "Chaco" continental low. Standard deviation of summer warm days at the weather stations in Argentina north of 40°S (right)

Predictability is defined as the measure to which the future states of a system can be predicted based on knowledge of the current and past state of the system. Research on predictability and their sources is a central part of creating new and improved forecasts. Advances in this area depend critically on observations and models. Typically, the predictability analysis of a phenomenon begins with theoretical considerations or empirical analysis based on observations (for example, an analysis of the lagged correlation between two or more variables). From this perspective, it is essential to have long records of multivariate observations for both predictors and potential predictors (National Academies of Sciences, Engineering, and Medicine 2016).

In terms of seasonal prediction, the predictability of the system resides mainly in the ocean because the ocean presents adjustment times slower than the atmosphere (Slingo and Palmer 2011) and in the soil moisture, since the lack of soil moisture strongly reduced latent cooling and thereby amplified the surface temperature anomalies (Fischer et al. 2007b). Many previous studies have found that sea surface temperature (SST) anomalies in the equatorial Pacific associated with the El Niño-Southern Oscillation (ENSO) are responsible for most interannual variability in the Southern Hemisphere (Vera et al. 2004; Garreaud et al. 2009, among others). A positive relationship between temperature anomalies and ENSO phases was found over a large part of South America, i.e., El Niño is associated with temperatures above normal while La Niña is associated with cold anomalies. More recently, Rusticucci et al. (2017) showed that the impact of El Niño events on extreme temperatures in Argentina presents monthly differences, favoring warming during the austral winter (more nights and warm days) and colder conditions in austral summer (fewer warm and more frequency of cold days). The opposite pattern occurs during La Niña: December, January, and February exhibited more frequency of warm days compared to its climatology, while December and February showed fewer cold days. However, on the climate of South America, the SST of the South Atlantic Ocean is also important. Rusticucci et al. (2003) showed that extreme temperature events in Argentina presented a greater seasonal correlation with the Atlantic than the Pacific, reflecting the importance of the "orographic barrier" of the Andes in the atmospheric circulation. The only exception refers to warm events in spring, for which the warming of the equatorial Pacific (the ENSO pattern) appears as the dominant mode.

Soil moisture is a key variable in the climate system due to its impacts on water and energy balances (Jaeger and Seneviratne 2011). Therefore, the relationship between soil moisture and mean and extreme temperature has been extensively studied using both observations and climate models (Whan et al. 2015 and references cited therein). Seneviratne et al. (2010) described the positive feedback between soil moisture-temperature: a decrease in soil moisture leads to a decrease in evapotranspiration (ET), this decrease in ET leads to an increase in sensible heat flux and then produces an increase in temperature, the increase in temperature leads to a higher-demand for evaporation and, therefore, to a potential increase in ET despite the dry conditions which possibly lead to a greater decrease in soil moisture. Several studies affirm that the soil moisture–temperature interaction increases the variability of summer temperatures: dry soils cause hot extreme frequencies increment (Oglesby and Erickson III 1989; Atlas et al. 1993; Seneviratne et al. 2006; Fischer et al. 2007a, 2007b; Lorenz et al. 2010; Hirschi et al. 2011; Mueller and Seneviratne 2012). Moreover, Jaeger and Seneviratne (2011) showed that the effect of soil moisture on temperature is asymmetric with greater impact on maximum temperature. Globally, Mueller and Seneviratne (2012) have found a strong relationship between the deficit of

precipitation and the subsequent occurrence of warm extremes of maximum temperature in several places in the world, such as large areas of North America, South America, Europe, Australia, and parts from China. They suggested that the prediction of warm days can be improved in these regions by incorporating soil moisture as a predictor. In particular for central and northern Argentina, these authors found significant negative correlations between the number of warm days in the warmest month of the year and the Standardized Precipitation Index (SPI) of 3, 6, and 9 months of the previous month.

The objective identification of temperature predictors together with their probabilities of occurrence is useful for the short- and medium-term forecasting as well as for future climate projections. However, seeking of temperature extreme predictors to incorporate them as independent variables in statistical models is a little-explored area in the world and particularly in Argentina. A study that explored a seasonal forecast in Argentina can be found at Solman and Núñez (1999). In that work, the authors estimated satisfactorily the mean values of winter and summer maximum and minimum temperature in the central region of Argentina using stepwise multiple regressions. Nevertheless, these authors conducted a forecast of mean minimum and maximum temperatures, and they did not perform a forecast of extreme temperature frequencies. Another useful statistical technique to predict time series when there are a large number of variables is to use principal component analysis followed by multiple linear regression (Çelik 2018). This combination is called principal component regression (PCR). This method reduces the complexity of the multidimensional system by maximizing the variance of component loadings and by eliminating the invalid components (Zhang et al. 2013). PCR has been employed to predict monthly precipitation in cost regions of Equator (de Guenni et al. 2016), seasonal rainfall in Southeast South America (Zamboni et al. 2010), heavy rainfall in Paraguay (Doss-Gollin et al. 2018), and rainfall variations in the Pacific Islands (Yu et al. 1997). Moreover, PCR has been widely used to predict rainfall over India during the monsoon season (Rajeevan et al. 2000; Nair et al. 2013), in the winter season (Nageswararao et al. 2016), and to predict the post-monsoon tropical cyclone activity (Biswas and Kundu 2018). Furthermore, it was used in air quality studies (Statheropoulos et al. 1998; Abdul-Wahab et al. 2005; Rajab et al. 2013; Tong et al. 2018) and to reconstruct different ENSO indices (Srivastava and Sinha Ray 2000; Barrett et al. 2018). However, this technique was not used to seasonal forecast extreme temperature in Argentina. Moreover, at present, Argentine Weather Service only provides a seasonal forecast for mean temperatures.

This study has two main objectives: first, to identify predictors of summer warm days by using lagged correlation analysis, and second, to fit a statistical model in order to predict frequencies of summer warm days in Argentina northern of 40°S. The lagged correlation is an important tool to study the relationship between two variables because one variable may have a delayed response to the other variable and, in that case, it can be used as a predictor of the other. Among all predictors analyzed, the persistence of warm days was especially considered by correlating summer warm days with those observed in the previous season (spring). We fitted a statistical model using PCR to the observed frequencies of summer warm days, and we evaluated the performance of the model over a test sub-sample. Furthermore, we explore potential predictors under ENSO neutral conditions which allowed us to have a more complete analysis for those years in which one of the main sources of predictability is not active.

2 Data and methodology

2.1 Extreme temperature index

Daily maximum temperatures of December, January, and February (DJF) were used in this work. These data were provided by the National Institute of Agricultural Technology (INTA, http://siga2.inta.gov.ar/#/data) and the Argentine National Meteorological Service (SMN, https://www.smn.gob.ar/) from 83 conventional weather stations located north of 40°S in Argentina in the 1970–2015 period. All the stations have less than 10% of missing data in the period analyzed.

The freely available R-Climdex software was used (at http://etccdi.pacificclimate.org/software.shtml, accessed in May 2015) in order to perform a quality control check of each station. Quality control is vital when working with extreme events, since an outlier may be due to errors in the measurement or transcription of the data or be truly an extreme value, so a detailed study of each one was carried out. In all cases, we verified that the maximum temperature was higher than the minimum temperature. Data with values far above or below the expected values were also checked. These aspects were analyzed individually, considering the information from nearby stations and temperature evolution on previous and subsequent days.

After analyzing the outliers, one extreme temperature index was calculated: warm days (TX90), defined as the percentage of days with maximum temperature above the 1981–2010 90th percentile, following the guidelines established by international literature on the subject based on recommendations by the WMO Joint Commission for Climatology (CCL) /CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI). The index was estimated on a monthly basis, and from them, seasonal values (December–February) were obtained. From now on, we will refer to these indices as DJF TX90. According to the definition of the index, the summer warm days present a mean value

of 10% in the climatological period (1981–2010). The interannual variability, measured by the standard deviations, range from 4.5 to 13% depending on the station, although most of the largest standard deviations are located in the center and northeastern Argentina (Fig. 1).

Since we employed data from 83 meteorological stations in Argentina north of 40°S, to display the results here, we grouped the meteorological stations in climatological homogeneous regions and selected one station of each group. To do this, we used the methodology Partitioning Around Medoids proposed by Kaufman and Rousseeuw (1990), or also known as K-medoids for its similarity with K-means. We considered eight different regions, according to the number of regions which are considered by the SMN to their seasonal forecast of mean temperature. The resulting groups are shown in Fig. 2.

2.2 Predictors data

The predictors' data were obtained from different data sources. Monthly atmospheric and oceanic indices were acquired from the website of the National Oceanic & Atmospheric Administration (NOAA) (http://www.esrl.noaa.gov/psd/data/climateindices/list/, accessed June 2016). The indices considered were Niño 1 + 2, Niño 3, Niño 3.4, Niño 4, Niño Modoki and Oceanic Niño Index (ONI), the Southern Oscillation Index (SOI), Tropical Southern Atlantic Index (TSA), Atlantic Multidecadal Oscillation (AMO), and Pacific Decadal Oscillation (PDO).

To represent other modes of atmospheric variability, different indices previously constructed by other authors were used. The Southern Annular Mode (SAM) index was defined as Marshall (2003), which considers a zonal average of sea-level pressure at 40°S and 65°S based on six stations close to each latitude (http://www.nerc-bas.ac.uk/icd/gjma/sam.html, consulted on 5/10/2016). In the Indian Ocean, the dipole mode index (IOD, °C) was defined by Saji et al. (1999) using the SST anomalies in the northwest (NW) regions (50–70°E, 10°S–10°N) and the southeast region (SE) (90–110°E, 10°S–0°N) of the equatorial Indian Ocean. The IOD index values were obtained from the website of the International Research Institute for Climate and Society (IRI) (http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCDC/.ERSST/.version4/.IOD /.C1961–2015/, accessed 06/21/2016).

In order to characterize the position and intensity of the subtropical jet, we considered the following indices: the zonal wind at 250 hPa (U250) averaged over 25–35°S and 50–70°W, zonal wind at 200 hPa averaged over 20–30°S and 45–65°W (U1), and over 30–40°S and 45–65°W (U2) and the quotient between both (U1/U2) defined by Rusticucci et al. (2017) and Barros et al. (2002). For the construction of all indices, reanalysis data from the National Centers for Atmospheric Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) Reanalysis 1 (Kalnay et al. 1996) was used because it covers the entire studied period (1970–2015) and it was previously employed in several studies in the region.

Data from the NCEP/NCAR Reanalysis 1 were also employed to evaluate the influence of the atmospheric blocking and the South Atlantic and Pacific anticyclones on the extremes of temperature indices. The presence of atmospheric blocking over the Pacific Ocean was represented by three indices which use the monthly zonal wind as the main variable: blocking index (IB) at 120°W (Rutllant and Aceituno 1991; Rutllant 2004), blocking index at 70°W

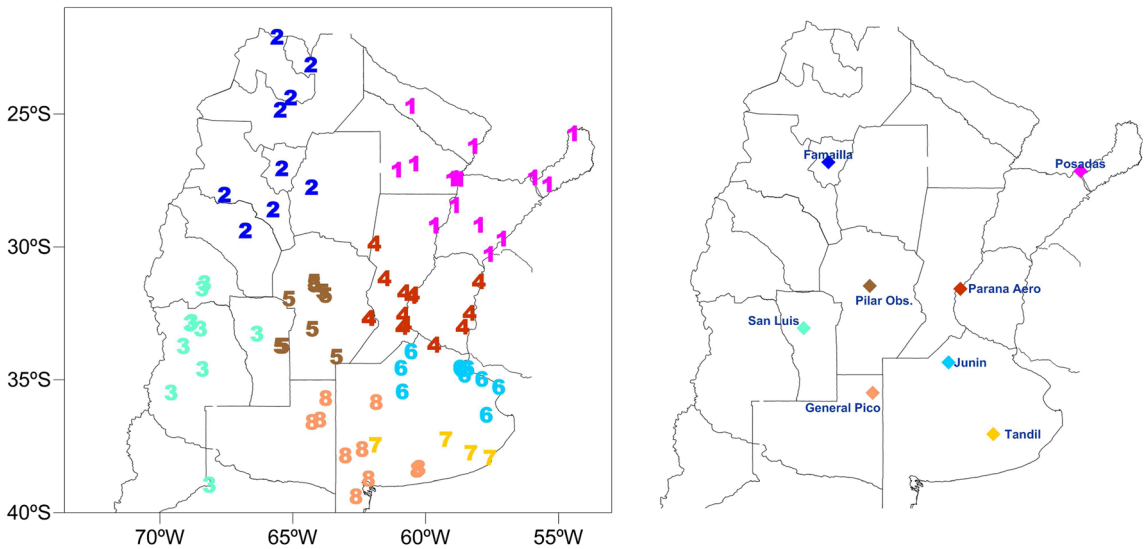

Fig. 2 Stations employed and clustering of them in eight regions (left): 1—northeastern, 2—northwestern, 3—cuyo, 4—south Litoral, 5—central, 6—northeastern Buenos Aires, 7—southeastern Buenos Aires, 8—central-south. Stations selected of each region (right)

(B70) and blocking index at 100°W (B100) defined by Alessandro (2014). Using geopotential height at 1000 hPa, we estimated six indices to characterize the intensity and position of the anticyclonic center (González et al. 2015): AAINT, AALAT, AALON for the SAA and APINT, APLAT, and APLON for the SPA. In addition, we estimated another index in the western region of the SAA (25–35°S, 30–45°W) by spatially averaging the monthly anomalies of geopotential height at 1000 hPa, identified from now by the initials ANOM (Rusticucci et al. 2017). The anomalies were estimated with respect to the base period 1981–2010. More details of these indices can be found in Table 1.

Warm air advection produced by the north wind is another important mechanism that affects surface temperature. The advection over northern and central Argentina was represented by the Z3-Z1 and V925 indices which reflect geopotential height differences at 850 hPa and meridional wind at 925 hPa, respectively (Table 1).

To represent soil moisture conditions, the SPI index was used as a proxy (Mueller and Seneviratne 2012). The SPI (McKee et al. 1993, 1995) is a powerful and flexible index that is simple to calculate. In fact, precipitation is the only required input parameter. The SPI was designed to quantify the precipitation deficit for multiple timescales. It quantifies observed precipitation as a standardized departure from a selected probability distribution function that models the raw precipitation data. The raw precipitation data are typically fitted to a gamma distribution and then transformed to a normal distribution. The SPI values can be interpreted as the number of standard deviations by which the observed anomaly deviates from the long-term mean. The SPI can be created for differing periods of 1-to-36 months, using monthly input data. Positive SPI values indicate greater than median precipitation, and negative values indicate less than median precipitation. Because the SPI is normalized, wetter and drier climates can be represented in the same way; thus, wet periods can also be monitored using the SPI. Employing monthly accumulated precipitation data from the same stations as those used to calculate DJF TX90, we estimated 1-, 3-, 6-, 9-, and 12-month SPI indices by using the software available on the website (http://drought.unl.edu/MonitoringTools/DownloadableSPIProgram.aspx). For those months that presented more than 10% of missing data, they were assigned a missing data code to the monthly accumulated precipitation. On short timescales, the SPI is closely related to soil moisture, while at longer timescales, the SPI can be related to groundwater and reservoir storage. Moreover, we calculated the first principal component of the SPI (PC1 SPI) in order to obtain a single monthly time series for each of the SPI indices which allow us to capture some characteristics of the spatial pattern of these variables.

In order to evaluate the influence of Atlantic SST on extreme temperature, we considered four indices. One of them represents the dipole of the Atlantic Ocean on the coasts of Brazil, the South Atlantic Subtropical Dipole index (SASD, Wainer et al. 2014), and three indices based in Barrucand et al. (2008): SST30, SST36, and SST46 which involve spatial averages of SST near the coast of Southern South America at 30°S, 36°S and 46°S.

Finally, to evaluate the effect of the South American monsoon, two monthly indices were considered based on the outgoing longwave radiation (OLR). OLR is a measure of the amount of energy emitted to space by earth's surface, oceans, and atmosphere. The OLR values are often used as a proxy for convection in tropical and subtropical regions since cloud top temperatures are indicative of cloud height (higher clouds producing lower values of OLR). In particular, OLR values below 240 Wm^{-2} are representative of convection (Kousky 1988). The first index used to represent the monsoon was the OLRI index (Wm^{-2}), which consists of the OLR anomaly in the area of the Amazon forest covered by (0–15°S; 45–75°W) (González et al. 2016). The second index was the South Atlantic Convergence Zone (SACZ) index defined as the spatial average of the OLR anomalies between 30 and 17°S and 45–20°W (Cerne and Vera 2010). Positive values of this index represent inhibited SACZ conditions while negative values indicate the presence of an active SACZ. The reference period for calculating the OLR anomalies for both indices was 1981–2010. The OLR dataset source was the NCEP/NCAR Reanalysis 1. These indices, as well as the other indices considered in this work, are summarized in Table 1.

2.3 Preliminary analysis: trends and wavelet transform

The best-fit linear trend is often used to describe the change in a climatological series (Vincent et al. 2005). Linear trends of DJF TX90 were calculated for the 1970–2015 period. Then, these trends were statistically tested by using the non-parametric Mann-Kendall test in R (McLeod 2011) at 5% of significance level. The purpose of the Mann-Kendall test (Mann 1945; Kendall 1975; Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time.

It is known that in the climate series may appear "jumps," periodic or quasi-periodic events, which do not necessarily remain in time but appear for a few years, and then disappear, or remain as weak signals in the system (Barrucand et al. 2008). Wavelet transform of a function is the improved version of Fourier transform. Fourier transform is a powerful tool for analyzing the components of a stationary signal, but it failed for analyzing the non-stationary signal (Sifuzzaman et al. 2009). The traditional Fourier transform uses sine and cosine base functions that have an infinite span and are

Table 1 Description of the indices considered as candidate for predictors

	Index name	Description	Reference
Modes of climate variability	Niño1 + 2	The Niño1 + 2 SST anomaly index is an indicator of far eastern tropical Pacific El Niño conditions, off the coasts of Peru and Chile. It is calculated with SSTs in the box 90–80°W, 10°S–0°.	https://www.esrl.noaa.gov/psd/data/climateindices/list/
	Niño3	The Niño3 SST anomaly index is an indicator of eastern tropical Pacific El Niño conditions. It is calculated with SSTs in the box 150–90°W, 5°S–5°N.	https://www.esrl.noaa.gov/psd/data/climateindices/list/
	Niño3.4	The Niño3.4 SST anomaly index is an indicator of central tropical Pacific El Niño conditions. It is calculated with SSTs in the box 170–120°W, 5°S–5°N.	https://www.esrl.noaa.gov/psd/data/climateindices/list/
	Niño4	The Niño4 SST anomaly index is an indicator of western tropical Pacific El Niño conditions. It is calculated with SSTs in the box 160–150°W, 5°S–5°N.	https://www.esrl.noaa.gov/psd/data/climateindices/list/
	SOI	The Southern Oscillation Index is a standardized index based on the observed sea level pressure differences between Tahiti and Darwin, Australia.	https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/
	ONI	3-month running mean of SST anomalies in the Niño 3.4 region (5°N–5°S, 120–170°W)	http://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php
	Modoki	El Niño Modoki Index = SSTA–0.5 SSTB–0.5 SSTC The three terms on the right-hand side of the equation are derived from the area-averaged SST anomaly over each of the regions A (195–140°W 10°S–10°N), B (110–70°W, 15°S–5°N), and C (235–215°W, 10°S–20°N), respectively.	Ashok et al. (2007)
	SAM	Difference between zonal mean of the sea level pressure at 40°S and 65°S based on 12 stations	Marshall (2003)
	IOD	Indicator of the east–west SST gradient across the tropical Indian Ocean. It is calculated as the difference of the Western Tropical Indian Ocean (50–70°E; 10°S–10°N) and Southeastern Tropical Indian Ocean (90–110°E; 10°S–0°)	Saji et al. (1999)
	TSA	Anomaly of the average of the monthly SST from Eq-20 S and 10 E–30 W. Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) and NOAA OI 1 × 1 datasets are used to create index. Climatology is 1971–2000.	Enfield et al. (1999)
	PDO	Leading PC of monthly SST anomalies in the North Pacific Ocean, poleward of 20°N	Zhang et al. (1997); Mantua et al. (1997)
	AMO	Compute the area-weighted average over the N Atlantic, basically 0 to 70°N. Then, this time series is detrended.	Enfield et al. (2001)
Regional atmospheric circulation	U250	Zonal wind at 250 hPa spatially averaged in the region between 25 and 35°S and 50–70°W. Represents the intensity of subtropical jet stream.	Rusticucci et al. (2017)
	U1	Zonal wind at 200 hPa averaged over 20–30°S and 45–65°W. Represents the intensity of the northern flank of the subtropical jet stream.	Barros et al. (2002)
	U2	Zonal wind at 200 hPa averaged over 30–40°S and 45–65°W. Represents the intensity of the southern flank of the subtropical jet stream	Barros et al. (2002)
	U1/U2	Ratio between the two previous indices. This index captures the position of the subtropical jet.	Barros et al. (2002)
	IB	Zonal wind at 200 hPa centered at 50°S and 120°W. IB = 0.5 * (U30 + U35 + U65 + U70 − U45 + U55–2 * U50) where the number beside U corresponds to the latitude. The IB reaches maximum positive values with the presence of a high block between 50°S and 70°S, and a trough between 50°S and 30°S.	Rutllant and Aceituno (1991); Rutllant (2004)
	B70	Zonal wind at 500 hPa at the longitude of 70°W. B70 = U(30°S) + U(60°S)-2 U(45°S) When a persistence blocking occurred, this index takes positive values	Alessandro (2014)
	B100	Idem B70 but at the longitude of 100°W	Alessandro (2014)

Table 1 (continued)

	Index name	Description	Reference
	AAINT	Anomalies of the intensity of the South Atlantic Anticyclone center estimated as the maximum of geopotential height at 1000 hPa in the region 0–40°S and 0–45°W. Climatology: 1981–2010 Positive values represent an intensification of the anticyclone.	González et al. (2015)
	AALAT	Anomalies of the latitude of the South Atlantic Anticyclone center. Climatology: 1981–2010 Positive values imply a shift to the north of its mean latitudinal position.	González et al. (2015)
	AALON	Anomalies of the longitude of the South Atlantic Anticyclone center. Climatology: 1981–2010. Positive values of AALON correspond to an eastward shift of its mean longitudinal position.	González et al. (2015)
	APINT	Anomalies of the intensity of the South Pacific Anticyclone center estimated as the maximum of geopotential height at 1000 hPa in the region 0–40°S and 72.5–150°W. Climatology: 1981–2010 Positive values represent an intensification of the anticyclone.	González et al. (2015)
	APLAT	Anomalies of the latitude of the South Pacific Anticyclone center. Climatology: 1981–2010. Positive values imply a shift to the north of its mean latitudinal position.	González et al. (2015)
	APLON	Anomalies of the longitude of the South Pacific Anticyclone center. Climatology: 1981–2010. Positive values of APLON correspond to an eastward shift of its mean longitudinal position.	González et al. (2015)
	ANOM	Anomalies of geopotential height at 1000 hPa in the western region of the South Atlantic Anticyclone (25–35°S, 30–45°W). Climatology: 1981–2010	Rusticucci et al. (2017)
	Z3-Z1	Gradient of geopotential height at 850 hPa. Z1 and Z3 are the geopotential heights averaged between 25 and 35°S at 65°W and between 20 and 30°S at 42.5°W, respectively	Barros et al. (2002)
	V925	Meridional wind at 925 hPa spatially averaged over 28–35°S and 57–64°W. This index intends to represent the low-level jet.	Rusticucci et al. (2017)
Atlantic SST and South American monsoon	SST30	Anomalies of sea surface temperature spatially averaged in the region 25.7–33.3°S and 45–50.6°W (climatology 1981–2010)	Barrucand et al. (2008)
	SST36	Anomalies of sea surface temperature spatially averaged in the region 33.3–39°S and 45–56.2°W (climatology 1981–2010)	Barrucand et al. (2008)
	SST46	Anomalies of sea surface temperature spatially averaged in the region 41–50.5°S and 58.1–65.6°W (climatology 1981–2010)	Barrucand et al. (2008)
	SASD	Represents the South Atlantic Subtropical Dipole index. It uses the difference between the spatial average of the SST anomalies in the NE Pole (15–25°S, 0–20°W) and in the SO Pole (30–40°S, 10–30°W)	Wainer et al. (2014)
	SACZ	Anomalies of outgoing long wave (OLR) spatially averaged in the region 17–30°S and 20–45°W. Climatology 1981–2010	Cerne and Vera (2010)
	OLRI	OLR anomalies in the area of the Amazon forest covered by (0–15°S; 45–75°W). Climatology 1981–2010	González et al. (2016)
Soil moisture conditions	SPI1	On short timescales, the SPI is closely related to soil moisture. A 1-month SPI map is very similar to a map displaying the percentage of normal precipitation for a 30-day period. In fact, the derived SPI is a more accurate representation of monthly precipitation because the distribution has been normalized.	McKee et al. (1993); Edwards and McKee (1997)
	SPI3	A 3-month SPI refle cts short- and medium-term moisture conditions and provides a seasonal estimation of precipitation. A 3-month SPI at the end of November compares the SON precipitation total in that particular year with the SON precipitation totals of all the years on record for that location.	McKee et al. (1993); Edwards and McKee (1997)
	SPI6		McKee et al. (1993);

Table 1 (continued)

Index name	Description	Reference
	The 6-month SPI indicates seasonal to medium-term trends in precipitation. A 6-month SPI at the end of November compares the precipitation total for the May–November period with all the past totals for that same period.	Edwards and McKee (1997)
SPI9	The 9-month SPI provides an indication of inter-seasonal precipitation patterns over medium timescale duration. This time period begins to bridge a short-term seasonal drought to those longer-term droughts that may become hydrological, or multi-year, in nature.	McKee et al. (1993); Edwards and McKee (1997)
SPI12	A 12-month SPI is a comparison of the precipitation for 12 consecutive months with the same 12 consecutive months during all the previous years of available data. The SPI at these time scales reflect long-term precipitation patterns.	McKee et al. (1993); Edwards and McKee (1997)
PC1 SPI1	First principal component obtained from the SPI1 index of the 83 stations.	
PC1 SPI3	First principal component obtained from the SPI3 index of the 83 stations.	
PC1 SPI6	First principal component obtained from the SPI6 index of the 83 stations.	
PC1 SPI9	First principal component obtained from the SPI9 index of the 83 stations.	
PC1 SPI12	First principal component obtained from the SPI12 index of the 83 stations.	

globally uniform in time (Pasquini et al. 2006). Wavelet transform is a method designed to optimize both time and frequency resolution by choosing the best window width for a particular frequency band (Kestin et al. 1998).

The global wavelet spectrum of a signal can be computed by averaging the local spectra over the temporal variable and allows determining the dominant periods of time series. Such plots may be compared with the spectrum derived from the Fourier transform since the Fourier spectrum and global wavelet transform represent the power over the entire signal (Kirby and Swain 2013). However, the comparison between the global wavelet power spectrum and the classical Fourier spectrum made by Perrier et al. (1995) shows that, at small scale, if the Fourier spectrum is too steep, the wavelet spectrum is strongly biased by the specific wavelet used in the analysis. At other scales, Percival (1995) shows that the global wavelet spectrum provides an unbiased and consistent estimation of the true power spectrum of a time series.

The wavelet transform is a very appropriate analysis tool for the study of non-stationary processes that involve multiple scales of variability. It has been used in many fields, such as image processing (Ren et al. 2017), optics (Buraga-Lefebvre et al. 2000), quantum mechanics (Frantziskonis and Deymier 2003), biometrics (Doghmane Doghmane et al. 2018), and geophysical series (Wei et al. 2018), among other fields. The wavelet transform was also widely applied in the meteorological field, some examples are for atmospheric macrocirculation indices, such as ENSO (Kestin et al. 1998) or North Atlantic Oscillation (Pozo-Vázquez et al. 2001); for boundary layer height (Pal et al. 2016); and for surface air temperature in South America (Kayano and Sansígolo 2009; Naumann and Vargas 2012). The presence of signals that appear for a few years and then disappear can hinder the forecast. This method allowed us to detect different periodicities in the DJF TX90 series and associated physical processes.

To apply the wavelet transformation to the DJF TX90 series, we used the WaveletComp package for R (Roesch and Schmidbauer 2014). This package calculates the wavelet power spectrum by applying the Morlet wavelet. The choice of the mother wave will depend on the type of study to be carried out, and the need to obtain a better temporal or frequency discrimination. In the case of geophysical series, it is widely accepted in the literature that the most suitable mother wave is the "Morlet" wave, a flat wave, modulated by a Gaussian (Foufoula-Georgiou and Kumar 1995; and the references therein). It is well known in signal processing that the Morlet mother wavelet gives the best time-frequency localization (Le 2017). Baliunas et al. (1997) have employed both Mexican hat and Morlet wavelets in a study of temperature data taken over three centuries in central England. The better temporal resolution of the Mexican hat is put to use in filtering trends in the time domain, whereas the superior frequency resolution of the Morlet wavelet is used for spectral studies

of the data. Considering that we were interested in having better frequency discrimination in order to assess the dominant periodicities of the series rather than a temporal localization, the Morlet wavelet was selected. In particular, the Morlet wave was employed by Barrucand et al. (2008) to analyze the main variability modes of extreme temperature indices, Elsanabary and Gan (2014) to study seasonal rainfall variability of the Upper Blue Nile Basin and sea surface temperatures, Zitto et al. (2016) to analyze temperatures and ice core oxygen isotope series at high latitudes of Southern Hemisphere, Zhou and Liu (2017) to study meteorological droughts, and Barrucand et al. (2018) to analyze climate indices, among many other studies.

Since time series have a finite length, errors will occur at the beginning and end of the wavelet power spectrum. The cone of influence is the region of the wavelet spectrum in which edge effects become important. More details could be found in Torrence and Compo (1998). Moreover, using this package, the trend of the series was filtered with the LOESS function.

To test the null hypothesis that a period is irrelevant in a given time, the package offers the option of performing simulations incorporating white noise into the data. Statistical significance tests for wavelet power spectra consist of deriving theoretical wavelet spectra for white noise processes and using these to establish significance levels and confidence intervals. In our case, we made 100 simulations. In case of missing data, they were replaced with the climatological values.

2.4 Predictability of time series

In order to find predictors of DJF TX90, we calculated 1-month lagged Spearman correlations for each meteorological station and all the indices mentioned in Sect. 2.2, i.e., DJF TX90 was correlated with the November predictors proposed in this study. We use only 1-month lag because several climate services consider 1-month lag to elaborate their seasonal forecast of precipitation and temperature, such as the International Research Institute or the Argentina National Meteorological Service. Nevertheless, some indices which describe processes over a longer timescale were included, such as SPI at different scales or the ONI index. Spearman correlation coefficient is a measure of non-parametric association between two continuous random variables that do not follow a normal distribution. The statistical significance of the correlation coefficient was tested using AS 89 algorithm considering a significance level of 10% (Best and Roberts 1975). The null hypothesis states that there is no monotonic association between the two variables in the population, against the alternative that there is a monotonic correlation present. The critical value of Spearman's rank correlation coefficient for 45 observations and a significance level of 10% for a two-tailed test is 0.248, according to Zar (1984) Table B.19. It is important to note that some stations have missing data; hence, this critical value could be a little different.

In the particular case of ONI, the index of September–November (SON) was correlated with DJF TX90. Moreover, to study the persistence of the series, we also calculated and tested the correlation of DJF TX90 with November TX90 and TX90 in the previous season, i.e., SON TX90. Seasonal linear trends are removed from the DJF TX90 time series to avoid a spurious empirical relationship.

For main predictors found, the mean value of DJF TX90 was estimated for those years in which each predictor registered an above-normal value (greater than the second tercile) and for those years in which the value of the index did not exceed the first tercile (below normal). Then, the difference between both means was calculated. The statistical significance was tested using a student t test for the mean of two samples with a confidence level of 95% and the non-parametric Wilcoxon rank sum test (Wilcoxon 1945). This test is equivalent to the Mann-Whitney U test, which evaluates the differences between frequency distributions. Similar results were obtained with student t and Wilcoxon test. In general, significant differences present absolute values above 4% of days. In addition, the spatial patterns of the significant correlations and those obtained for the DJF TX90 differences were also similar.

Another aspect to be considered is the multicollinearity between predictors. Multicollinearity exists whenever an independent variable is highly correlated with one or more of the other independent variables in a multiple regression equation (Allen 1997). This fact can be a problem for some statistical models—such as multiple linear regression—because it can increase the variance of the estimated coefficients and make the estimation very sensitive to minor changes in the model. The result is that the coefficients are unstable and difficult to interpret. Therefore, to detect collinearity, we analyzed the simultaneously Spearman correlations between predictors, having previously filtered the trends, and their significance was tested using AS 89 algorithm (Best and Roberts 1975). The correlations with absolute values above 0.3 were significant at 5%.

Finally, we studied the predictability of DJF TX90 under ENSO neutral conditions. As we mentioned before, the ENSO is one of the predominant sources of interannual climate variability in the world (Trenberth and Caron 2000). The extreme phases of the ENSO, El Niño and La Niña, modulate the mean temperature and precipitation, as well as the frequency of occurrence of extremes of these variables. In this sense, Osman and Vera (2016) found that the predictability of temperature, precipitation, and mean circulation over South America increases slightly in ENSO years. Therefore, under ENSO neutral conditions, one of the main drivers of the atmospheric circulation is inactive, which makes the seasonal prediction more complicated. In order to find predictors of

DJF TX90 under ENSO neutral conditions, the indices and temperature extreme indices were correlated but considering only neutral years. To determine the neutral conditions, we considered the NOAA classification of the ENSO phases based on the ONI index. Seventeen neutral years occurred during the study period; hence, the critical value of Spearman correlation with a significance level of 10% is 0.41.

2.4.1 Modeling of DJF TX90 using PCR

If we have a large number of correlated variables, we could try to create a new set of explanatory variables which are linear combinations of the original set and which are orthogonal to each other (Butler and McNertney 1991). Principal components (PCs) are based on this idea. The PCR is similar to the standard multiple linear regression, but instead of using the predictors directly, the PCs of the predictors are used as independent variables. One of the main uses of PCR is to overcome the problem of multicollinearity that arises when two or more of the explanatory variables are close to being collinear. However, if one of the objectives is to identify a smaller set of predictors, PCR cannot be used for this purpose since PCs are, by definition, linear combinations of all the original predictors (Bunea et al. 2011). Because PCs depend on the scale of the variables, they were standardized.

By definition, predicting the future leads us into uncharted territory. The simplest way for us to quantify the ability of a predictive model to perform future data is to try to simulate this eventuality. Although we cannot, literally, gain access to the future before it happens, we can reserve some of our currently available data and treat it as if it were future data (Steinberg 2014). Therefore, to estimate how well the model can predict new values, it is required to separate a dataset that was not previously used to build the model. For this reason, DJF TX90 was divided into a training sub-sample, or also called a calibration sub-sample and a test sub-sample. After having fitted the PCR in the training data subset, the model can be evaluated in the test dataset to obtain a final objective idea of how the models might work in the not seen data (Brownlee 2016). We calculated bias, root mean square error (RMSE), mean absolute error (MAE), and Spearman correlation coefficient between observation in the test sample and the modeled values (Al-Lami et al. 2017). Bias can convey useful information but should be interpreted cautiously because positive and negative errors will cancel out. The RMSE and MAE solve this problem: RMSE represents the sample standard deviation of the differences between predicted values and observed values, while MAE is a linear score which means that all the individual differences are weighted equally in the average. So, RMSE penalizes the higher difference more than MAE. On the other hand, the correlation coefficient is a measure of the strength of the association between the observed and predicted values.

The 1970–1999 period was considered for the training sub-sample and the 2000–2015 period for the test sub-sample. It is desirable that the calibration period be longer than the test period to increase the reliability of the statistical analysis and allow a range of natural variability, assuming that the expected changes in the average climate would be within the range of natural variability (von Storch et al. 1993). The PCs were generated using all the indices in Table 1. In particular the PC1 SPI indices reflect spatial characteristics of soil moisture over the whole studied area. In PCR, the number of principal components is typically chosen by cross-validation in the calibration sub-sample (James et al. 2013). We compute ten-fold cross-validation error for each possible value of the number of principal components used, and then we chose the number of principal components associated with the lowest adjusted RMSE.

In the validation period, we estimate the bootstrapping prediction intervals at 90% confidence level. A prediction interval is a quantification of the uncertainty on a prediction. The uncertainty comes from the errors in the model itself and noise in the input data. If we only produce deterministic forecasts, there is no way of telling how accurate the forecasts are. However, if we also produce prediction intervals, then it is clear how much uncertainty is associated with each forecast. Following Hastie et al. (2010) and Kumar and Srivastava (2012), we performed the steps mentioned below:

1. Randomly resample the training dataset with replacement, each sample with the same size as the original training set.
2. Fit the PCR model to the resample data in the training period
3. Select the model with the number of principal components which produce the lowest cross-validation error.
4. Resample the residuals obtained for the selected model in step 3.
5. Predict DJF TX90 in the testing period using the model obtained in step 3.
6. Add the resample residuals to the predicted values.
7. Repeat all the previous steps 500 times in order to obtain many possible predictions.
8. Compute prediction intervals by calculating percentiles 5 and 95 to the forecast series obtained in step 7.

3 Results

3.1 Preliminary analysis of DJF TX90 series: trends and wavelet analysis

First, we analyzed the trends of DJF TX90 in the 1970–2015 period and their statistical significance (Fig. 3). We

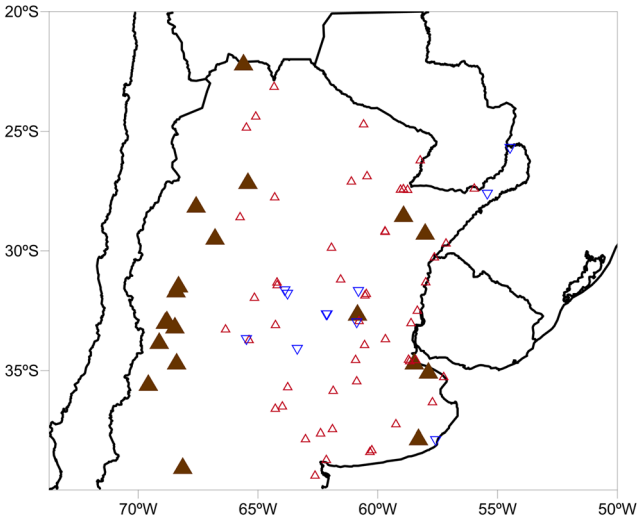

Fig. 3 Trends of DJF TX90 in the 1970–2015 period [%/year]. Significant positive trends are represented with solid up triangles, non-significant positive trends with small up triangles, and non-significant negative trends with down triangles. The range of significant positive trends at 5% is 0.12–0.48%of days/year

found a predominance of positive trends, in accordance with global warming. The significant positive trends are located mainly in western Argentina. Only in the center of the country, some negative tendencies were observed, although they were not significant. The trends were filtered for the subsequent analysis.

In order to detect non-stationary periodicities, we applied wavelet transform to the series. Figure 4 shows the wavelet power spectrum for one station of each region shown in Fig. 2. Fourier periods are shown on the vertical axis and the years on the horizontal axis. The white contours indicate the significant periods at 10% and the black lines the local maximums of the wavelet transform that provides an estimate of instantaneous periods. To the right of the power spectrum, we show the global wavelet power spectrum for each period: on the vertical axis, the Fourier periods are shown and, on the horizontal axis, the averages are presented. The points indicate those averages that were significant at 5% contrasting with the spectrum that would be obtained for the white noise.

At the eight stations, we appreciate a periodicity of 2–4 years, probably associated with the influence of ENSO over extreme temperature. This assumption will be deeply analyzed in the following section. Low-frequency modes are present, but these modes are discontinued in time. For example, in Famailla, San Luis, and General Pico, a 6–8-year periodicity was found in the 1980–1995 period approximately. In contrast, in Posadas and Tandil, an 8-year periodicity is presented in the first years of the period. These particularities might be a problem at the moment of modeling these series. Finally, in San Luis and Junin, a weak periodicity of 16 years was detected in the first half of the period. This wavelet analysis was useful to notice that processes with a decadal variability could also be a predictor of extreme temperature. Note that the results outside the cone of influence were not taken into account in the analysis since errors will occur at the beginning and ends of the spectrum because of the limited time series.

3.2 Predictors of DJF TX90

3.2.1 TX90 as predictor

First, in order to evaluate the persistence, we studied the correlation between November TX90 and DJF TX90, and the correlation between SON TX90 and DJF TX90. The results are presented in Fig. 5. Reflecting persistence, both maps show significant positive correlations in a large number of stations. A positive correlation implies that the current state of the index endures for the next season, i.e., if SON TX90 presents a low value, DJF TX90 will also show a low value.

Therefore, the previous values of TX90 are relevant for predicting the summer occurrence of warm days, and it could be considered a predictor of this variable. However, we found differences between the two correlation maps. The correlation between SON TX90 and DJF TX90 shows a higher (lower) number of stations with significant correlations in the north (central) Argentina in comparison with the correlation between November TX90 and DJF TX90.

3.2.2 Modes of climate variability

The main modes of climate variability that affect the circulation of South America through teleconnections were correlated with DJF TX90. Figure 6a presents the percentage of stations with significant correlations at 10%. The results showed that the ENSO and PDO in November presented a strong association with DJF TX90 at several stations. Among ENSO indices, Niño 4 had the best performance with 64% of stations significantly correlated.

Figures 7a and b show the difference between mean DJF TX90 for years when November Niño 4 exceeded the second tercile and mean DJF TX90 for years when the indices were lower than the first tercile, and the same for PDO. Both fields present negative values in the entire domain, except in northwestern Argentina. For Niño 4, the significant negative differences are located mainly in the center and southeastern of the region, while PDO shows the maximum absolute values in northeastern Argentina. According to these results, El Niño condition in November is associated with fewer warm days in summer, and the opposite occurs under La Niña events. The link between the ENSO and the extreme temperature is probably due to the changes generated by the ENSO on circulation and precipitation in Argentina. Finally, other modes, such as SAM or IOD, would not have an influence on DJF TX90 while November TSA was significantly correlated with DJF TX90 only at a few stations. These results are in agreement

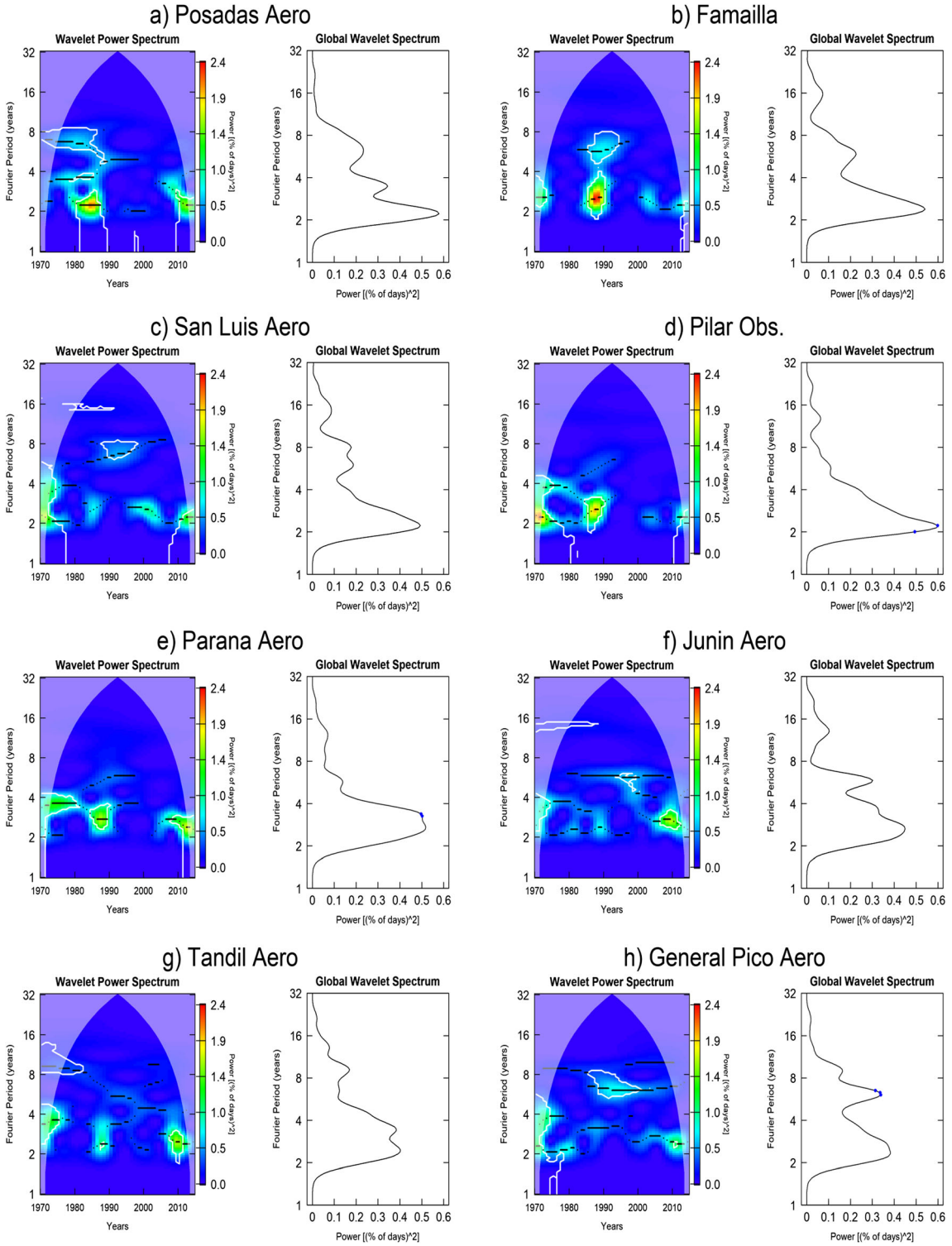

Fig. 4 Wavelet of DJF TX90 power spectra (shaded, units: % of days^2). Wavelet power spectrum: Fourier periods (in years) are shown on the vertical axis and the calendar years on the horizontal axis. White contours indicate the significant periods at 10%, and the black lines the local maximums of the wavelet transform that provides an estimate of instantaneous periods. Global wavelet spectrum: on the vertical axis the Fourier periods (in years) are shown and on the horizontal axis, the averages of the wavelet power are presented (right, units: % of days^2). The points indicate those averages that were significant at 5% contrasting with the spectrum that would be obtained for the white noise. **a** Posadas Aero (Group 1). **b** Famailla (Group 2). **c** San Luis Aero (Group 3). **d** Pilar Obs. (Group 4). **e** Parana Aero (Group 5). **f** Junin Aero (Group 6). **g** Tandil Aero (Group 7). **h** General Pico Aero (Group 8)

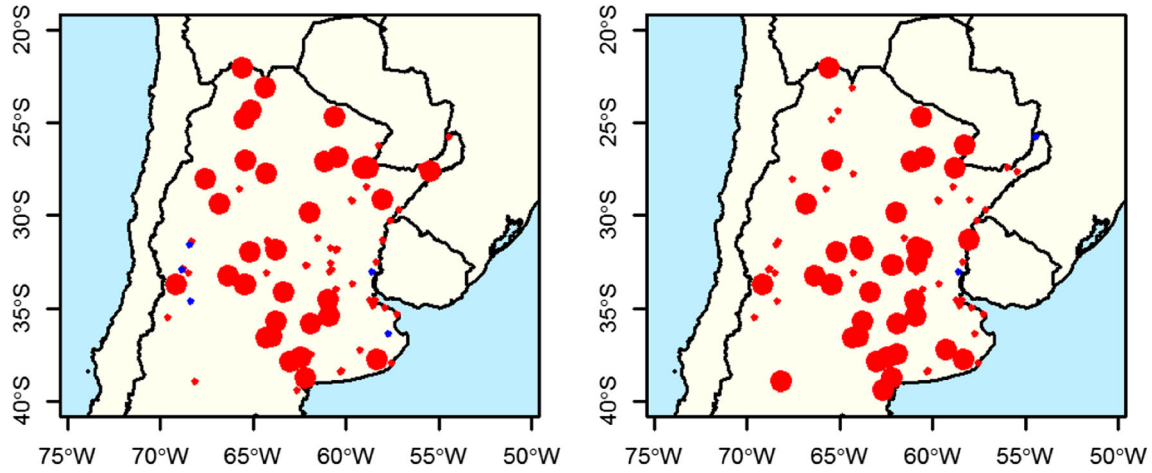

Fig. 5 Correlation between DJF TX90 and November TX90 (left) and SON TX90 (right). Significant positive correlations (big dots), non-significant correlations (small dots). Significant level 5%; the general lower limit of significance correlation at each station is 0.3

with previous work that found that the influence of SAM and IOD is particularly strong during the spring. In particular, the positive phase of the SAM is associated with an intensification of an anticyclonic anomaly of high levels, a weakening of the convergence of humidity and a decrease in precipitation over southeastern South America (SESA) and warmer conditions in Patagonia (Silvestri and Vera 2003; Gillett et al. 2006). Regarding the IOD, Saji et al. (2005) found significant partial correlations between IOD and mean temperature during SON in northern Argentina, Uruguay, Paraguay, Bolivia, and Brazil.

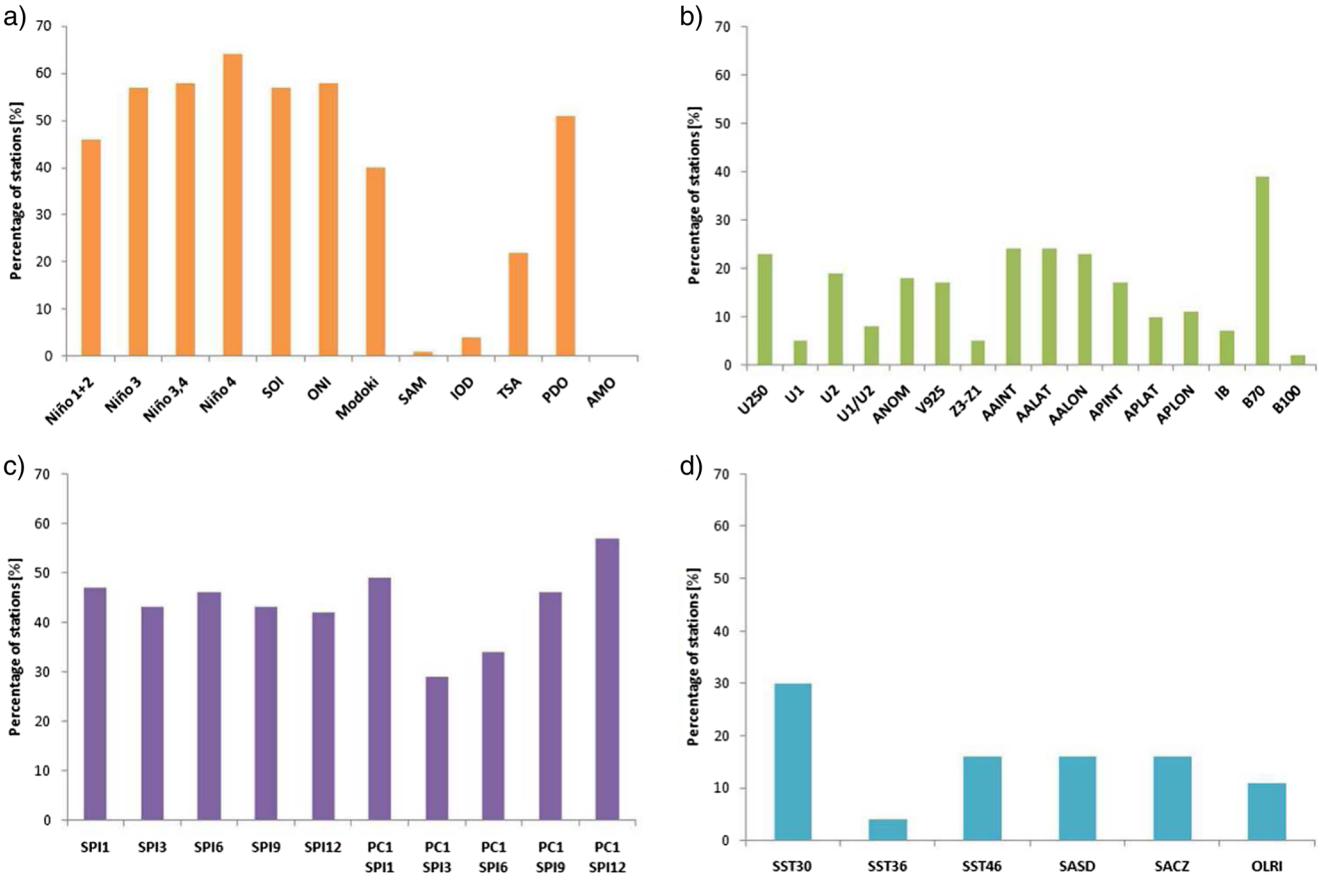

Fig. 6 Percentage of stations significantly correlated at 10%. The correlations were calculated between the different candidates to predictors in each November and subsequent DJF TX90 in the 1970–2015 period. **a** Climate indices. **b** Regional atmospheric circulation. **c** Standardized precipitation indices. **d** Sea surface temperature and outgoing longwave radiation

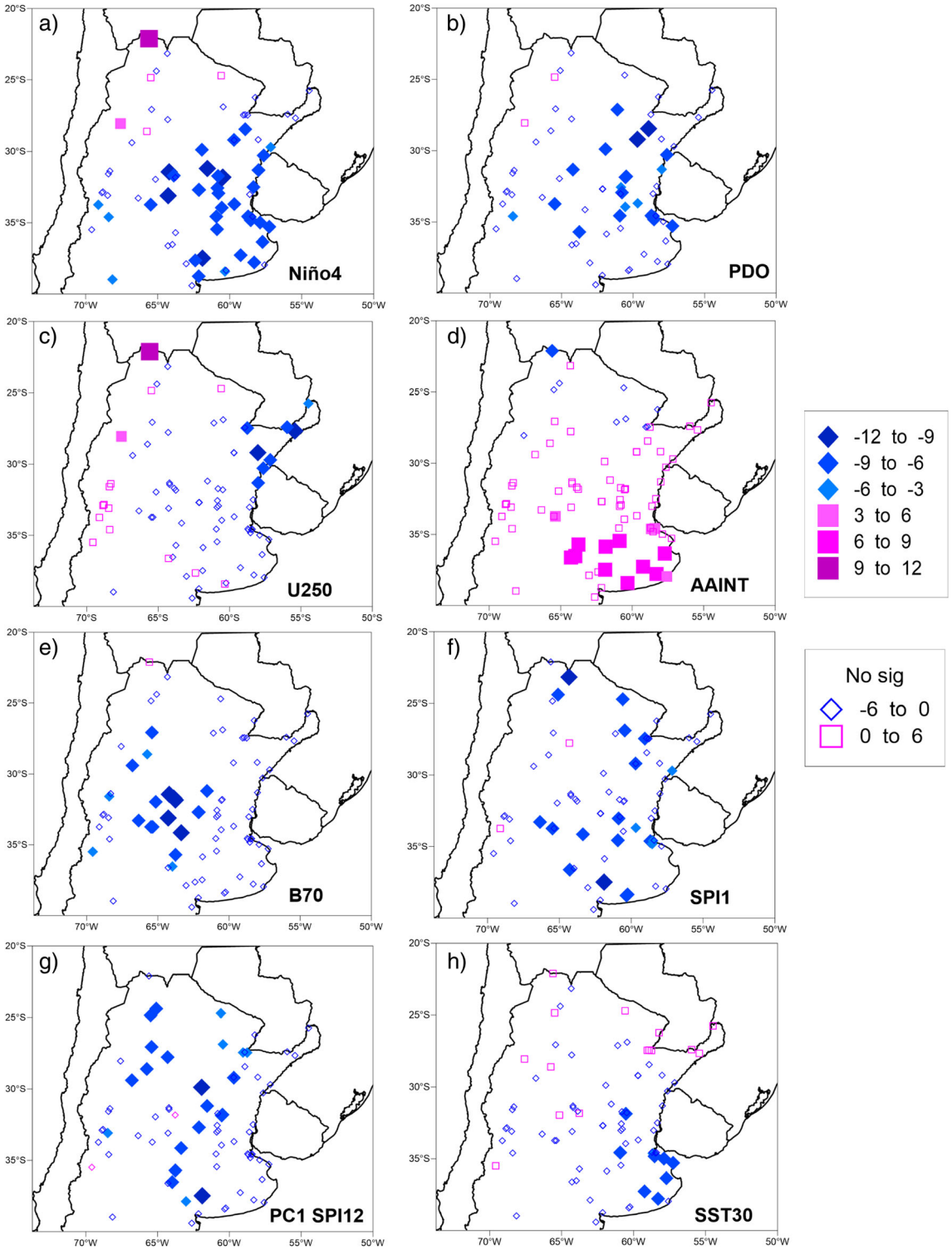

Fig. 7 Difference fields between DJF TX90 for years when the predictors exceeded the second tercile and the same for years when the index was lower than the first tercile [% of days] for **a** El Niño 4, **b** PDO, **c** U250, **d** AAINT, **e** B70, **f** SPI1, **g** PC1 SPI12 and **h** SST30. Significant level 5%

3.2.3 Regional atmospheric circulation

In general, the regional atmospheric circulation indices of November appear to have little and located influence in summer warm days (Fig. 6b). Some selected indices were analyzed: U250 (representative of the subtropical jet), AAINT (associated with SAA), and B70 (related to atmospheric blocking). We observed that an above-normal

intensity of the subtropical jet in November is associated with a lower (higher) occurrence of warm days in northeastern (northwestern) Argentina (Fig. 7c). Consistently, Barros et al. (2002) found similar results for mean temperature in northeastern Argentina since they observed that an intense jet is associated with more passages of cold fronts which produce a decrease of temperature.

In southern and southeastern parts of the studied region, AAINT might be a good predictor of warm days: this index is positively correlated with DJF TX90 (Fig. 7d). This relationship is probably due to the fact that an intense SAA favors an increase of warm advection from the north (Fig. S1). While in the central region, the occurrence of atmospheric blockings in November (high values of B70) is associated with below-normal occurrence of summer warm days; hence, B70 could be considered a predictor of DJF TX90 in that region (Fig. 7e).

3.2.4 Atlantic SST and South American monsoon

The Atlantic SST and convection associated with South American monsoon do not seem to be suitable predictors of DJF TX90. The most relevant predictor is SST30 which is significantly correlated with DJF TX90 in 30% of the weather stations analyzed (Fig. 6d). Significant negative differences were found in southeastern of the domain (Fig. 7h). Therefore, above-normal SST on the coast of southern Brazil in November is associated with a reduction of the number of warm days in summer. A warmer Atlantic on the coasts of Brazil generates more humidity available in the atmosphere that is advected to Argentina by the northern winds on the western flank of SAA and produces an increase of cloudiness and precipitation in eastern Argentina (González et al. 2016). Figure S2 endorses this hypothesis by showing a positive simultaneous correlation between SST30 and precipitation of the Global Precipitation Climatology Project (GPCP) in November. The link between previous rainfall and the extreme temperature is through soil moisture–atmosphere feedback: wet soils tend to inhibit extreme warm events because of evaporative cooling.

We did not find significant correlations between November SACZ and DJF TX90 at 5% significance level (critical value of significant correlations is 0.294), and only 16% of stations were significantly correlated at 10%. These few significant correlations may be due to the fact that convection in tropical regions operates more on intra-seasonal scales than on a seasonal scale (Cerne and Vera 2010; Alvarez et al. 2016).

3.2.5 Soil moisture conditions

The SPI was used as a proxy of soil moisture. For each meteorological station, we calculated the correlations between DJF TX90 and November SPI. We considered here the SPI computed for 1–3-, 6-, 9-, and 12-month accumulation periods. These correlations were statistically tested at a 10% significance level. The DJF TX90 index shows a strong correlation (between −0.25 and −0.6) with SPI in almost half of the locations (Fig. 6c). In addition, the correlations between DJF TX90 and the first principal component (PC1) of SPI indices were also significant at many stations. When considering the PC1 of SPI indices, the percentage of stations significantly correlated only improves considerably for SPI12.

As we mentioned before, one reason for the increase of summer warm days is the precipitation deficits that occur in the preceding months. The SPI allows us to represent the lack of rainfall during a prolonged period. Moreover, these precipitation deficits are usually the results of several climate drivers some of which were taken into account individually in this work (e.g., La Niña events produce a reduction of precipitation in southeastern South America) but only with 1-month lag. Consequently, the SPI can integrate the effect of several processes which influence the precipitation rate at multiple timescales and have a more complete analysis.

Figure 7f illustrates this relationship: when dry conditions (negative SPI1) are present in November, warm conditions in summer will probably be favored. Two sectors showed significant differences of DJF TX90 between opposite phases of soil moisture (wet and dry conditions): one located in the north and the other in central-south of the domain. This finding is consistent with Mueller and Seneviratne (2012) since they found a significant relationship between the SPI indices and the number of warm days in Argentina during the summer.

Furthermore, Fig. 7g indicates a negative correlation between November PC1 of SPI12 and DJF TX90, being significant in all central and northern regions. Therefore, soil moisture not only influences locally but a general dry condition benefits the occurrence of summer warm days.

3.3 Collinearity of predictors

The collinearity between the indices was analyzed in the 1970–2015 period through Spearman's simultaneous correlations, having filtered the trends previously (Fig. 8). As it was expected, the variables which represent the same phenomenon are significantly correlated, such as the ENSO indices, the first principal component of the SPI, the wind indices at upper levels, or the blocking indices. However, not only the predictors associated with the same phenomenon are correlated with each other. For example, ENSO indices are significantly correlated at 5% with IOD (0.57–0.65), subtropical jet (0.29–0.56), PDO (0.4–0.65), the first principal component of SPI of 1, 3, and 6 months (0.39–0.61), among others, which emphasized how relevant is ENSO for seasonal predictability of extreme

temperature in Argentina. In brackets, we display the range of the absolute values of the significant correlations.

In particular, analyzing only the predictors in Fig. 7, we found that El Niño4 index is significantly correlated at 5% with PDO (0.46) and SST30 (0.50), and the U250 index is associated with B70 (0.46). Correlation values are in brackets. Then, we fitted a multiple regression model to each station in the training period by using only these eight predictors and calculated the variance inflation factor (VIF) to measure the multicollinearity (see more details in the supporting information). The VIFs take values between 1.9 and 4 which indicate some multicollinearity, but not extremely severe. However, we decided to choose a model which was capable of dealing with this problem in order to obtain better results, since multicollinearity makes the estimation very sensitive to minor changes in the model. Potential solutions include remove some highly correlated predictors or use partial least squares regression (PLS) or principal components regression methods that cut the number of predictors to a smaller set of uncorrelated components (Jackson 1991).

3.4 Modeling DJF TX90

To solve the collinearity problem of predictors, we fitted a PCR to all DJF TX90 series on the training sub-samples and then we applied this model to the test sub-samples in the 2000–2015 period. Figure 9 shows the observed and modeled DJF TX90 for eight representative stations, one from each region. In addition, Table 2 presents statistical scores to evaluate the performance of the simulation using PCR. The correlations between the observed and modeled DJF TX90 were significant at 5%, except in Famailla. The critical value of Spearman's rank correlation coefficient for 15 values and a significance level of 5% for a two-tailed test is 0.52 according to Zar (1984) Table B.19. The prediction interval gives an interval within which we expect DJF TX90 to lie with a specified probability, in this case, 90%. Figure 9 shows that the prediction intervals usually contain the observed values with only a few exceptions.

Considering all statistical scores, Famailla station in northwestern Argentina (group 2) showed the worst results of the model, with a RMSE of 11.0% and MAE of 8.4%, while in other regions, the RMSE and MAE were approximate of 7.0% and 5.0% respectively. At the eight stations, the RMSE of predictions were of the order of one standard deviation of DJF TX90 obtained in the climatological period. These poor results for the stations in northwestern Argentina were possibly related to that we found few predictors for DJF TX90: only PC1 SPI12 and ENSO showed significant correlations at several stations (Fig. S3). On the other hand, Paraná Aero in group 5 was one of the best according to the scores: it had the lowest RMSE and MAE, a bias of only − 0.6% and a correlation of 0.70.

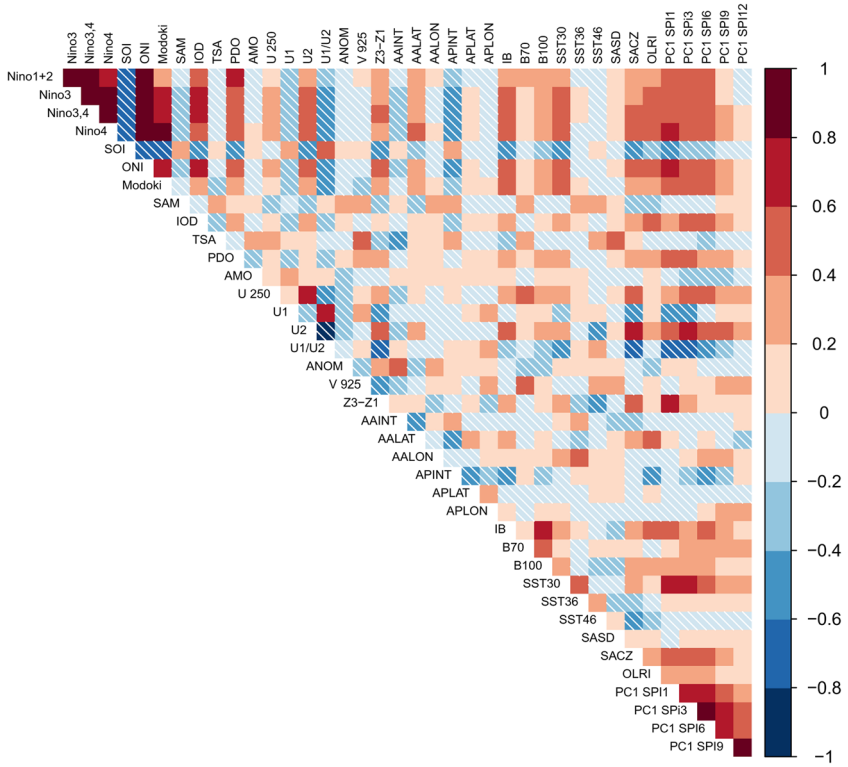

Fig. 8 Simultaneous correlations between the indices in November. Negative values are presented with stripes. Significant correlations at 5% present values above 0.3

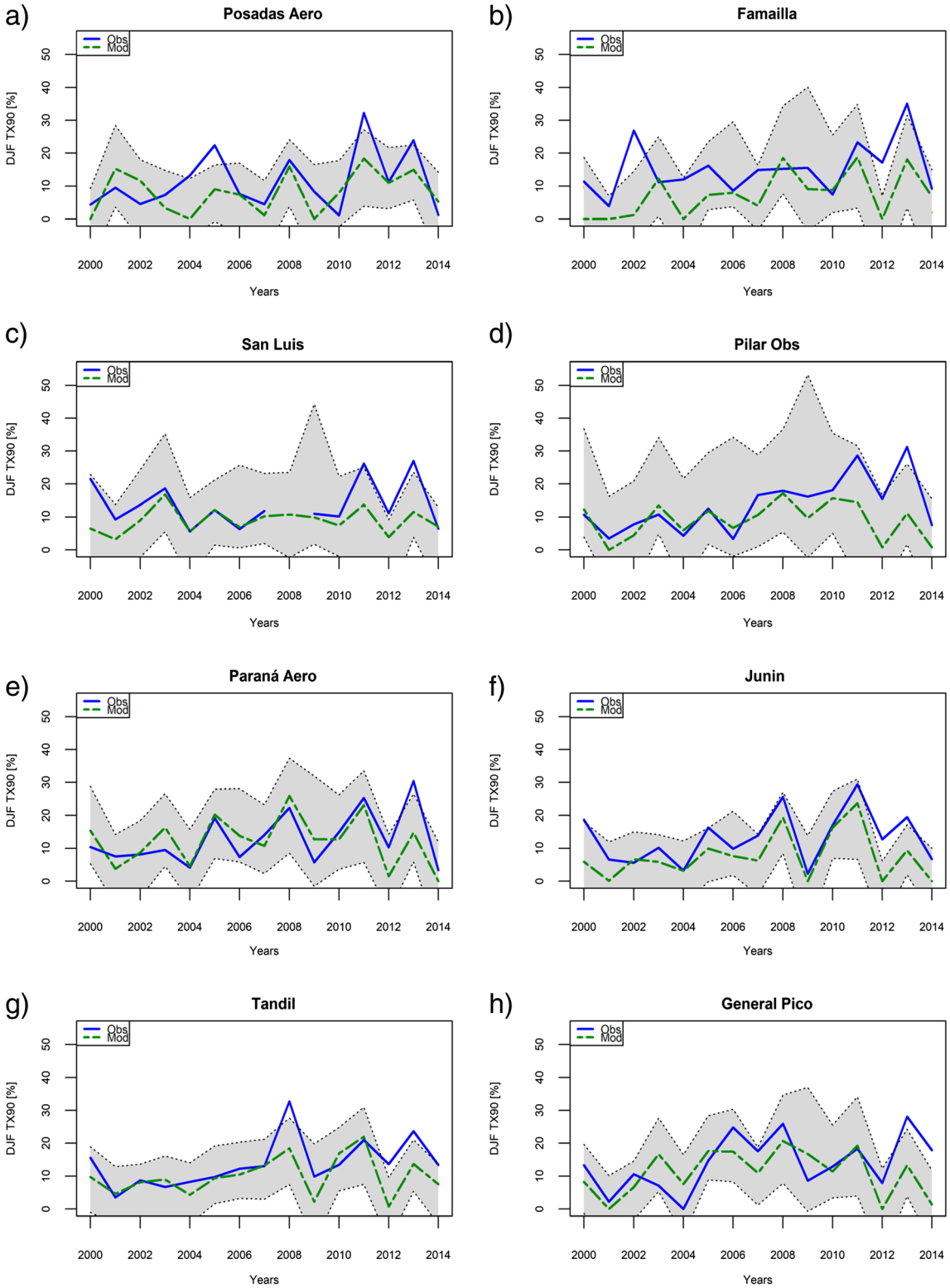

Fig. 9 Observed (solid line) and modeled (dashed line) DJF TX90 using PCR in the test sub-sample for eight stations, one from each region. The prediction intervals at 90% of confidence level is shown in gray

In the eight locations analyzed, we found negative biases of the model, i.e., the model underestimated the DJF TX90. In particular, the year 2013 was clearly underestimated by the model at the eight stations; even more, the observed value exceeded the upper limit of the prediction interval. Many regions of the country had their warmest December on record in

Table 2 Statistical scores (bias, RMSE, MAE, correlation coefficient) calculated between observed and modeled DJF TX90 using PCR in the test sub-sample

	Bias [% of days]	RMSE [% of days]	MAE [% of days]	Correlation
Posadas—Group1	− 3.2	7.7	6.4	0.53
Famailla—Group 2	− 7.7	11.0	8.4	0.26
San Luis—Group 3	− 4.8	7.3	5.0	0.63
Pilar Obs—Group 4	− 4.6	8.2	5.9	0.66
Paraná Aero—Group 5	− 0.6	6.0	4.7	0.70
Junín—Group 6	− 5.6	6.9	5.7	0.74
Tandil—Group 7	− 3.7	6.5	4.8	0.60
General Pico—Group 8	− 2.8	8.0	6.7	0.56

2013 due to the most significant heat wave in Argentina since February 1987 (NOAA 2014). Probably, the occurrence of this unusual extreme heat wave made it difficult for the model to represent it properly. The analysis of the atmospheric dynamics shows that the high temperatures during December 2013 were primarily associated with an intensification of the SACZ, which jointly caused simultaneous extreme rainfall events in Southeastern Brazil, and the presence of an anticyclonic circulation anomaly over central Argentina (Hannart et al. 2015).

3.5 Predictors of DJF TX90 under ENSO neutral conditions

We showed that ENSO is an important source of predictability for DJF TX90. In addition, we noticed that the year 2013, in which the model had problems to simulate the high values of DJF TX90, was a neutral year. These reasons motivated to study what happens when the strong signal of the ENSO is off and if the predictors under ENSO neutral conditions are different from those found in Sect. 3.2.

Under ENSO neutral conditions, we observed that the significant correlations of DJF TX90 with almost all the indices are reduced (Fig. 10). Both PDO and SPI lose relevance as predictors of DJF TX90, while B70 does not show big changes in the percentage of stations with significant correlations. This behavior is probably due to the fact that ENSO is strongly correlated with PDO and SPI, while B70 is independent of ENSO conditions (Fig. 8). Among the predictors that become more relevant in neutral conditions can be mentioned AMO, Z3-Z1, U1, and IOD.

In northeastern Argentina, PDO and SPI indices are still the main predictors, but a new predictor is added: AALON, which presents significant positive correlation at several stations of the region (Fig. S4a). Therefore, a shift to the east of the SAA favors an increase of DJF TX90 under ENSO neutral conditions. In northwestern Argentina, AMO and Z3-Z1 become predictors of DJF TX90 at some stations, while SPI indices continued to be prominent predictors (Fig. S4b). In the Cuyo region, the main predictors are U1, Z3-Z1, and PC1 SPI12 (Fig. S4c). In the central region, PDO is no longer a useful predictor under neutral conditions, but PC1 SPI12 remains as a predictor at a few stations and Z3-Z1 is incorporated as a new one (Fig. S4d). In region 5, the PC1 SPI of 1, 9, and 12 months are useful predictors of DJF TX90 at almost all stations. Other predictors are PDO, U250, and Z3-Z1 (Fig. S4e). In region 6, there are few significant correlations, and the predictors present a great variability among stations (Fig. S4f). In the southeast of the domain, AAINT is an important predictor under all phases of ENSO, while SPI is only important at one station (Fig. S4g). Finally, in region 8, the main predictors are U1, U1/U2, APINT, and SPI12 (Fig. S4h).

Therefore, we noticed that the geopotential height gradient at 850 hPa (Z3-Z1) becomes a relevant predictor of DJF TX90 in several regions under ENSO neutral conditions. Both variables are negatively correlated, so a strong gradient inhibit the occurrence of summer warm days in neutral phases probably because it favors the entrance of humid air from the north that increases the rainfall in the region. The zonal wind in the northern flank (U1) is positively correlated with DJF TX90 in the Cuyo and south regions. In this regard, Zamboni et al. (2010) have shown evidence of a simultaneous linear relationship at interannual time scales between precipitation anomalies over southeastern South America and the interannual variability of upper-level wind during spring, summer, and fall. Therefore, stronger winds at upper levels are associated with rainfalls in northeastern Argentina, and clear skies in the south and western region which could explain the increase in summer warm days.

4 Summary and discussion

In the present work, we made a detailed study about the seasonal predictability of extremely warm daily summer temperatures (DJF TX90) in Argentina north of 40°S, using data in the 1970–2015 period. The predictability of an event is derived from many processes and mechanisms that exhibit a wide range of time scales. Therefore, we analyzed a large number of predictors: global climate indices, regional circulation indices, standard precipitation indices, and SST in the nearby ocean.

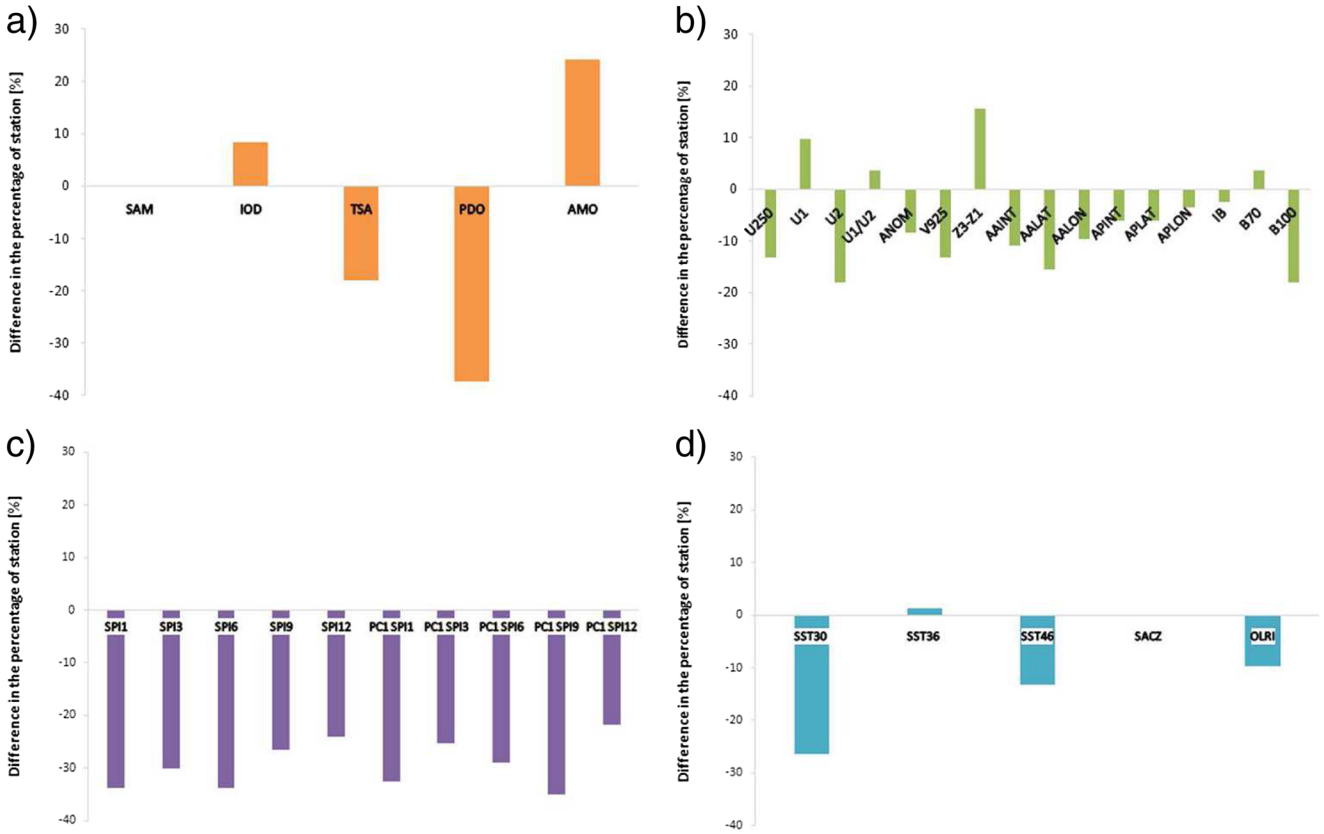

Fig. 10 Difference of the percentage of stations correlated significantly (at 10%) with the indices when only the ENSO neutral conditions are considered with respect to the previously observed in Fig. 6

We detected significant positive trends in DJF TX90 series, mainly in the western part of the domain, which was filtered for the rest of the analysis. Wavelet transform allows us to distinguish the main periodicities and their persistence over time. We found periodicities of 2–4 years, probably associated with ENSO, and low-frequency signals (approximately 8 years) that were not present during the entire study period.

Using Spearman lagged correlations, we looked for predictors of summer warm days. The TX90 index in the previous month and season are predictors of summer warm days which indicates persistence in the series. Other main predictors are ENSO which is reinforced by PDO, SPI in different terms, and atmospheric blocking at 70°W (B70). The occurrence of El Niño events is associated with a reduction of the summer warm days, mainly because El Niño produced an increase in precipitation in southeastern South America and a reduction in northwestern Argentina. By teleconnection processes, there are significant impacts on seasonal and monthly precipitation amounts in several regions of South America during the different phases of the ENSO (e.g., Ropelewski and Halpert 1987, 1989; Aceituno 1988; Rao and Hada 1990; Grimm et al. 1998, 2000; Vargas et al. 1999; Grimm 2003, 2004). This increase in the rainfall is associated with an increment of cloudiness, and a reduction of sensible heat flux joined with an increase of the latent heat flux, which could explain the reduction of summer warm days during El Niño. PDO is generally associated with an influence on climate similar to that produced by the ENSO in South America, but with much less influence (Garreaud et al. 2009). In this sense, Kayano and Andreoli (2007) indicated that the ENSO and the PDO act constructively on the precipitation of South America when both are in the same phase. For extreme temperature, we also found that ENSO and PDO act constructively on summer warm days, a positive phase of PDO in November is associated with lower values of DJF TX90 just like El Niño.

Due to the robustness of their thermohydrodynamical structure, blocking episodes tend to prevent the normal progression of transient systems (e.g., cold fronts and extratropical cyclones), favoring adverse and persistent meteorological conditions over the nearby regions (Mendes et al. 2008). The negative correlations found between DJF TX90 and B70 were consistent with Alessandro (2014) who found negative anomalies of mean temperature and an increase of precipitation northern of 38°S in summer.

Both short-term and long-term soil memory appear to be important predictors of summer warm days: a low value of SPI indicates dry conditions which benefit warmer conditions due to soil moisture–atmosphere feedback. This relationship between the precipitation deficit and the subsequent occurrence of hot extremes is due to the fact that in dry soils, the

evaporation cooling is reduced, and the sensible heat flow is increased (Seneviratne et al. 2010; Hirschi et al. 2011; Mueller and Seneviratne 2012). Therefore, moist soils act over temperature through feedbacks by reducing the occurrence of warm extremes of maximum temperature. We found that SPI indices are also relevant predictors of DJF TX90.

We decided to fit the principal component regression to the DJF TX90, with all the predictors we had studied, in order to obtain a statistical model to predict extreme warm temperatures on a seasonal scale in summer. We employed this model because is capable of dealing with the collinearity of predictors. The performance of the PCR model applied in this work is reasonably good for the test period: the statistical scores showed significant correlations between the observed and modeled data at the eight stations analyzed with the exception of Famailla located in northwestern Argentina. The model presented negative bias at the eight locations, varying from −0.6% in Paraná to −7.7% in northwestern Argentina. In particular, we also noticed that at all the stations, the year 2013 was the most difficult to predict due to the unusual extreme heat wave that occurred in December, under an ENSO neutral phase.

Considering the relevance of ENSO as a source of predictability of DJF TX90, we decided to look for predictors under ENSO neutral conditions. We found more dispersed results between stations. The influence of SPI, principally for a 12-month accumulation period, is still important as predictor due to dry soils benefitting the occurrence of summer hot days, but new predictors appear, mainly associated with SAA intensity and position, the north advection, IOD and AMO. As the ENSO, the IOD modifies the general circulation via anomalous convection that produces Rossby waves (Saji and Yamagata 2003) and influences the climate in several regions around the world. Regarding the AMO, the atmospheric patterns associated with the positive phase include cyclonic anomalies over the Atlantic between 20°S and 50°N, wind speed reduction over the tropical Atlantic and increase in precipitation in the eastern tropical Atlantic, with opposite conditions during the negative phase of the AMO. Moreover, the correlation between the AMO and temperature anomalies is positive over a large part of the globe between 40°S and 50°N (Alexander et al. 2014). On the other hand, when the low-level jet (associated with high values of Z3-Z1 index) is strong, the moisture flux is strongest towards southeast South America increasing precipitation in this region, and when the low-level jet is weak, the moisture flux intensifies towards southeast Brazil increasing precipitation in the SACZ region (Doyle and Barros 2002; Liebmann et al. 2004; Marengo et al. 2004).

This article presents a statistical model approach to predict an extreme temperature index on a seasonal scale that can be useful to reduce the impacts of these extreme events on different sectors such as agriculture, health, and energy. The fitted PCR model is one of the first attempts to construct a seasonal forecast of an extreme temperature index in Argentina based on empirical methods. The satisfactory results obtained in this work encourage us to continue working to develop an operational probabilistic seasonal forecast based on this statistical model along with others that may be generated in the future.

Acknowledgments This research was supported by CONICET PIP 0137-Res 4248/16 and UBACyT 2018 20020170100357BA. We want to thank the National Weather Service of Argentina and National Institute of Agricultural Technology for providing the data for this study. The authors want to especially thank Dr. Mariela Sued and Dr. Ana Bianco for her collaboration.

References

Abdul-Wahab SA, Bakheit CS, Al-Alawi SM (2005) Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. Environ Model Softw 20:1263–1271. https://doi.org/10.1016/j.envsoft.2004.09.001

Aceituno P (1988) On the functioning of the Southern Oscillation in the South American sector. Part I: Surface climate. Mon Weather Rev 116:505–524

Agosta EA, Compagnucci RH (2008) The 1976/77 austral summer climate transition effects on the atmospheric circulation and climate in southern South America. J Clim 21(17):4365–4383. https://doi.org/10.1175/2008jcli2137.1

Alessandro AP (2014) Incidence and trend of blocking action situations on the temperature and precipitation in Argentina. Atmosfera 27(2): 141–163. https://doi.org/10.1016/S0187-6236(14)71106-3

Alexander MA, Kilbourne KH, Nye JA (2014) Climate variability during warm and cold phases of the Atlantic Multidecadal Oscillation (AMO) 1871–2008. J Mar Syst 113:14–26. https://doi.org/10.1016/j.jmarsys.2013.07.017

Al-lami AM, AM AL-S, YK AL-T (2017) Parameterization of the downward long wave radiation under clear-sky condition in Baghdad, Iraq. Asian J Appl Sci 10:10–17. https://doi.org/10.3923/ajaps.2017.10.17

Allen MP (1997) The problem of multicollinearity. Understanding regression analysis. Springer, Boston, MA. https://doi.org/10.1007/978-0-585-25657-3_37

Alvarez MS, Vera CS, Kiladis GN, Liebmann B (2016) Influence of the Madden Julian Oscillation on precipitation and surface air temperature in South America. Clim Dyn 46:245–262. https://doi.org/10.1007/s00382-015-2581-6

Ashok K, Behera SK, Rao SA, Weng H, Yamagata T (2007) El Niño Modoki and its possible teleconnection. J Geophys Res 112: C11007. https://doi.org/10.1029/2006JC003798

Atlas R, Wolfson N, Terry J (1993) The effect of SST and soil moisture anomalies on GLA model simulations of the 1988 U.S. summer drought. J Clim 6:2034–2048. https://doi.org/10.1175/1520-0442(1993)006<2034:TEOSAS>2.0.CO;2

Baliunas S, Frick P, Sokoloff D, Soon W (1997) Time scales and trends in the Central England temperature data (1959–1990): a wavelet analysis. Geophys Res Lett 24(11):1351–1354. https://doi.org/10.1029/97GL01184

Barrett HG, Jones JM, Bigg GR (2018) Reconstructing El Niño Southern Oscillation using data from ships' logbooks, 1815–1854. Part I: methodology and evaluation. Clim Dyn 50:845–862. https://doi.org/10.1007/s00382-017-3644-7

Barros V, Grimm A, Doyle M (2002) Relationship between temperature and circulation in Southeastern South America and its influence from El Nino and La Nina Events. J Meteorol Soc Jpn 80(1):21–32

Barrucand M, Rusticucci M, Vargas W (2008) Temperature extremes in the south of South America in relation to Atlantic Ocean Surface temperature and Southern Hemisphere circulation. J Geophys Res Atmos 113:D20111. https://doi.org/10.1029/2007JD009026

Barrucand MG, Zitto ME, Piotrkowski R, Canziani P, O'Neill A (2018) Historical SAM index time series: linear and nonlinear analysis. Int J Climatol 38:1091–1106. https://doi.org/10.1002/joc.5435

Best DJ, Roberts DE (1975) Algorithm AS 89: the upper tail probabilities of Spearman's Rho. J Royal Stat Soc, Series C (Applied Statistics) 24(3):377–379. https://doi.org/10.2307/2347111

Biswas HR, Kundu PK (2018) A principal component analysis based model to predict post-monsoon tropical cyclone activity in the Bay of Bengal using oceanic Niño index and dipole mode index. Int J Climatol 38:2415–2422. https://doi.org/10.1002/joc.5344

Brownlee J (2016) Master machine learning algorithms. https://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms/. Accessed 22 January 2018.

Bunea F, She Y, Ombao H, Gongvatana A, Devlin K, Cohen R (2011) Penalized least squares regression methods and applications to neuroimaging. NeuroImage 55(4):1519–1527. https://doi.org/10.1016/j.neuroimage.2010.12.028

Buraga-Lefebvre C, Coëtmellec S, Lebrun D, Özkul C (2000) Application of wavelet transform to hologram analysis: three-dimensional location of particles. Opt Lasers Eng 33(6):409–421. https://doi.org/10.1016/s0143-8166(00)00050-6

Butler MR, McNertney EM (1991) Estimating educational production functions: the problem of multicollinearity. Soc Sci J 28(4):489–499, ISSN 0362–3319. https://doi.org/10.1016/0362-3319(91)90026-Z

Çelik K (2018) Predicting chlorophyll-a concentrations in two temperate reservoirs with different trophic states using principal component regression (PCR). Oceanol Hydrobiol Stud 47(1):1–9. https://doi.org/10.1515/ohs-2018-0001

Cerne B, Vera CS (2010) Influence of the intraseasonal variability on heat waves in subtropical South America. Clim Dyn 36:2265–2277. https://doi.org/10.1007/s00382-010-0812-4

de Guenni LB, García M, Muñoz ÁG, Santos JL, Cedeño A, Perugachi C, Castillo J (2016) Predicting monthly precipitation along coastal Ecuador: ENSO and transfer function models. Theor Appl Climatol 129:1059–1073. https://doi.org/10.1007/s00704-016-1828-4

Doghmane H, Bourouba H, Messaoudi K, Bouridane A (2018) Palmprint recognition based on discriminant multiscale representation. J Electron Imaging 27(5):053032. https://doi.org/10.1117/1.JEI.27.5.053032

Doss-Gollin J, Muñoz X, Mason S, Pastén M (2018) Heavy rainfall in Paraguay during the 2015–2016 austral summer: causes and subseasonal-to-seasonal predictive skill. J Clim 31:6669–6685. https://doi.org/10.1175/JCLI-D-17-0805.1

Doyle ME, Barros VR (2002) Midsummer low-level circulation and precipitation in subtropical South America and related sea surface temperature anomalies in the South Atlantic. J Clim 15:3394–3411. https://doi.org/10.1175/1520-0442(2002)015<3394:MLLCAP>2.0.CO;2

Edwards DC, McKee TB (1997) Characteristics of 20th century drought in the United States at multiple time scales. Climatology Report Number 97–2. Colorado State University, Fort Collins.

Elsanabary MH, Gan TY (2014) Wavelet analysis of seasonal rainfall variability of the upper blue Nile basin, its teleconnection to global sea surface temperature, and its forecasting by an artificial neural network. Mon Weather Rev 142(5):1771–1791. https://doi.org/10.1175/mwr-d-13-00085.1

Enfield DB, Mestas AM, Mayer DA, Cid-Serrano L (1999) How ubiquitous is the dipole relationship in tropical Atlantic sea surface temperatures? J Geophys Res Oceans 104:7841–7848. https://doi.org/10.1029/1998JC900109

Enfield DB, Mestas-Nunez AM, Trimble PJ (2001) The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental U.S. Geophys Res Lett 28:2077–2080. https://doi.org/10.1029/2000GL012745

Fischer EM, Seneviratne S, Lüthi D, Schär C (2007a) Contribution of land–atmosphere coupling to recent European summer heat waves. Geophys Res Lett 34:L06707. https://doi.org/10.1029/2006GL027992

Fischer EM, Seneviratne S, Vidale P, Lüthi D, Schär C (2007b) Soil moisture–atmosphere interactions during the 2003 European summer heat wave. J Clim 20:5081–5099. https://doi.org/10.1175/JCLI4288.1

Foufoula-Georgiou E, Kumar P (1995) Wavelets in geophysics. Academic Press, 373 pp

Frantziskonis G, Deymier P (2003) Wavelet-based spatial and temporal multiscaling: bridging the atomistic and continuum space and time scales. Phys Rev B 68(2):024105. https://doi.org/10.1103/physrevb.68.024105

Garreaud RD, Vuille M, Compagnucci R, Marengo J (2009) Present-day South American Climate. Palaeogeogr Palaeoclimatol Palaeoecol 281:180–195. https://doi.org/10.1016/j.palaeo.2007.10.032

Gilbert RO (1987) Statistical methods for environmental pollution monitoring. Wiley, NY

Gillett NP, Kell TD, Jones PD (2006) Regional climate impacts of the southern annular mode. Geophys Res Lett 33:L23704. https://doi.org/10.1029/2006GL027721

González MH, Garbarini E, Romero P (2015) Rainfall patterns and the relation to atmospheric circulation in northern Patagonia (Argentina). Adv Environ Res 41:85–100

González MH, Garbarini E, Rolla AL, Eslamian S (2016) Meteorological drought indices: rainfall prediction in Argentina. In: Handbook of Drought and Water Scarcity. Vol. 1, Principle of Drought and Water Scarcity, Chapter 29, 540–567, Taylor & Francis Publishing (CRC Group) Editor: Saeid Eslamian. ISBN: 9781498731089 1498731082. Reino Unido, Abingdon.

Grimm AM (2003) The El Niño impact on summer monsoon in Brazil: regional processes versus remote influences. J Clim 16:263–280

Grimm AM (2004) How do La Niña events disturb the summer monsoon system in Brazil? Clim Dyn 22:123–138

Grimm AM, Ferraz SET, Gomes J (1998) Precipitation anomalies in southern Brazil associated with El Niño and La Niña events. J Clim 11:2863–2880

Grimm AM, Barros VR, Doyle ME (2000) Climate variability in southern South America associated with El Niño and La Niña events. J Clim 13:35–58

Hannart A, Vera CS, Otto FEL, Cerne B (2015) Causal influence of anthropogenic forcings on the Argentinian heat wave of December 2013. Bull Amer Meteor Soc 96(12):S41–S45. https://doi.org/10.1175/bams-d-15-00137.1

Hastie T, Tibshirani R, Friedman J (2010) The elements of statistical learning, second edition: data mining, inference, and prediction. Springer Series in Statistics. Retrieved from http://www.worldcat.org/isbn/0387848576

Hirschi M, Seneviratne SI, Alexandrov V, Boberg F, Boroneant C, Christensen OB, Formayer H, Orlowsky B, Stepanek P (2011) Observational evidence for soil moisture impact on hot extremes in southeastern Europe. Nat Geosci 4:17–21. https://doi.org/10.1038/ngeo1032

IPCC (2012) Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change - Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner GK, Allen SK, Tignor M, Midgley PM (Eds.) Available from Cambridge University Press, The Edinburgh Building, Shaftesbury Road, Cambridge CB2 8RU ENGLAND, 582 pp.

Jackson E (1991) A user's guide to principal components. Wiley, New York

Jaeger EB, Seneviratne SI (2011) Impact of soil moisture–atmosphere coupling on European climate extremes and trends in a regional climate model. Clim Dyn 36:1919–1939. https://doi.org/10.1007/s00382-010-0780-8}

James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning with applications in R. Springer, New York

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. Bull Amer Meteor Soc 77: 437–472. 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kaufman L, Rousseeuw PJ (1990) Finding groups in data: an introduction to cluster analysis. Wiley, New York

Kayano MT, Andreoli RV (2007) Relations of South American summer rainfall interannual variations with the Pacific Decadal Oscillation. Int J Climatol 27:531–540. https://doi.org/10.1002/joc.1417

Kayano MT, Sansígolo C (2009) Interannual to decadal variations of precipitation and daily maximum and daily minimum temperatures in southern Brazil. Theor Appl Climatol 97:81–90. https://doi.org/10.1007/s00704-008-0050-4

Kendall MG (1975) Rank Correlation Methods, 4th edn. Charles Griffin, London

Kestin TS, Karoly DJ, Yano J, Rayner NA (1998) Time–frequency variability of ENSO and stochastic simulations. J Clim 11: 2258–2272. https://doi.org/10.1175/1520-0442(1998)011<2258:TFVOEA>2.0.CO;2

Kirby JF, Swain CJ (2013) Power spectral estimates using two-dimensional Morlet-fan wavelets with emphasis on the long wavelengths: jackknife errors, bandwidth resolution and orthogonality properties. Geophys J Int 194(1):78–99. https://doi.org/10.1093/gji/ggt103

Kousky VE (1988) Pentad outgoing longwave radiation climatology for the South American sector. Rev bras meteorol 3:217–231

Kumar S, Srivastava A (2012) Bootstrap prediction intervals in non-parametric regression with applications to anomaly detection. Conference Paper. The 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining; 12–16 Aug. 2012; Beijing; China. https://ntrs.nasa.gov/search.jsp?R=20130014367

Le T (2017) Use of the Morlet mother wavelet in the frequency-scale domain decomposition technique for the modal identification of ambient vibration responses. Mech Syst Signal Process 95:488–505. https://doi.org/10.1016/j.ymssp.2017.03.045

Liebmann B, Kiladis GN, Vera CS, Saulo AC, Carvalho LMV (2004) Subseasonal variations of rainfall in South America in the vicinity of the low-level jet east of the Andes and comparison to those in the South Atlantic convergence zone. J Clim 17:3829–3842. https://doi.org/10.1175/1520-0442(2004)017<3829:SVORIS>2.0.CO;2

Lorenz R, Jaeger EB, Seneviratne SI (2010) Persistence of heat waves and its link to soil moisture memory. Geophys Res Lett 37(9): L09703. https://doi.org/10.1029/2010GL042764

Mann HB (1945) Non-parametric tests against trend. Econometrica 13: 163–171

Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC (1997) A Pacific interdecadal climate oscillation with impacts on salmon production. Bull Amer Meteor Soc 78:1069–1079. https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2

Marengo JA, Soares WR, Saulo C, Nicolini M (2004) Climatology of the low-level jet east of the Andes as derived from the NCEP–NCAR reanalyses: characteristics and temporal variability. J Clim 17:2261–2280. https://doi.org/10.1175/1520-0442(2004)017\2261:COTLJE[2.0.CO;2

Marshall G (2003) Trends in the Southern Annular Mode from observations and reanalyses. J Clim 16:4134–4143. https://doi.org/10.1175/1520-0442(2003)0162.0.CO;2

McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scale. In: Proceedings of the Eighth Conference on Applied Climatology, Anaheim, California, 17–22 January 1993. American Meteorological Society, Boston, pp 179–184

McKee TB, Doesken NJ, Kleist J (1995) Drought monitoring with multiple time scales. Ninth Conference on Applied Climatology, American Meteorological Society, Jan 15–20, 1995, Dallas TX, pp.233–236.

McLeod AI (2011) Kendall: Kendall rank correlation and Mann-Kendall trend test. R package version 2:2 https://CRAN.R-project.org/package=Kendall

Mendes MCD, Trigo RM, Cavalcanti IFA, DaCamara CC (2008) Blocking episodes in the Southern Hemisphere: impact on the climate of adjacent continental areas. Pure Appl Geophys 165(9–10): 1941–1962. https://doi.org/10.1007/s00024-008-0409-4

Mueller B, Seneviratne SI (2012) Hot days induced by precipitation deficits at the global scale. Proc Natl Acad Sci 109(31):12398–12403. https://doi.org/10.1073/pnas.1204330109

Nageswararao MM, Mohanty UC, Osuri KK, Ramakrishna SSVS (2016) Prediction of winter precipitation over northwest India using ocean heat fluxes. Clim Dyn 47:2253–2271. https://doi.org/10.1007/s00382-015-2962-x

Nair A, Mohanty UC, Acharya N (2013) Monthly prediction of rainfall over India and its homogenous zones during monsoon season: a supervised principal component regression approach on general circulation model products. Theor Appl Climatol 111:327–339. https://doi.org/10.1007/s00704-012-0660-8

National Academies of Sciences, Engineering, and Medicine (2016) Next generation earth system prediction: strategies for subseasonal to seasonal forecasts. The National Academies Press, Washington DC. https://doi.org/10.17226/21873

Naumann G, Vargas WM (2012) A study of intraseasonal temperature variability in southeastern South America. J Clim 25:5892–5903. https://doi.org/10.1175/JCLI-D-11-00482.1

NOAA National Centers for Environmental Information, State of the Climate: Global Climate Report for Annual 2013, published online January 2014, retrieved on May 31, 2018 from https://www.ncdc.noaa.gov/sotc/global/201313. Accessed 19 June 2018.

Oglesby RJ, Erickson III DJ (1989) Soil moisture and the persistence of North American drought. J Clim 2: 1362–1380. doi:10.1175/1520-0442(1989)002,1362: SMATPO.2.0.CO;2.

Osman M, Vera CS (2016) Climate predictability and prediction skill on seasonal time scales over South America from CHFP models. Clim Dyn 49:2365–2383. https://doi.org/10.1007/s00382-016-3444-5

Pal S, De Wekker SF, Emmitt GD (2016) Investigation of the spatial cariability of the convective boundary layer heights over an isolated mountain: cases from the MATERHORN-2012 experiment. J Appl Meteorol Climatol 55:1927–1952. https://doi.org/10.1175/JAMC-D-15-0277.1

Pasquini AI, Lecomte KL, Piovano EL, Depetris PJ (2006) Recent rainfall and runoff variability in central Argentina. Quat Int 158(1):127–139. https://doi.org/10.1016/j.quaint.2006.05.021.

Percival DP (1995) On estimation of the wavelet variance. Biometrika 82: 619–631

Perrier V, Philipovitch T, Basdevant C (1995) Wavelet spectra compared to Fourier spectra. J Math Phys 36:1506–1519

Pozo-Vázquez D, Esteban-Parra M, Rodrigo F, Castro-Díez Y (2001) A study of NAO variability and its possible non-linear influences on

European surface temperature. Clim Dyn 17:701–715. https://doi.org/10.1007/s003820000137
Rajab JM, MatJafri MZ, Lim HS (2013) Combining multiple regression and principal component analysis for accurate predictions for column ozone in Peninsular Malaysia. Atmos Environ 71:36–43. https://doi.org/10.1016/j.atmosenv.2013.01.019
Rajeevan M, Guhathakurta P, Thapliyal V (2000) New models for long range forecasts of summer monsoon rain fall over northwest and peninsular India. Meteorol Atmos Phys 73:211–225. https://doi.org/10.1007/s007030050074
Rao VB, Hada K (1990) Characteristics of rainfall over Brazil: annual variations and connections with the Southern Oscillation. Theor Appl Climatol 42:81–90
Ren R, Gu L, Fu H, Sun C (2017) Super-resolution algorithm based on sparse representation and wavelet preprocessing for remote sensing imagery. J Appl Remote Sens 11(2):026014. https://doi.org/10.1117/1.JRS.11.026014
Roesch A, Schmidbauer H (2014) WaveletComp: computational wavelet analysis. R package version 1.0. https://CRAN.R-project.org/package=WaveletComp
Ropelewski CF, Halpert MS (1987) Global and regional scale precipitation patterns associated with the El Niño–Southern Oscillation. Mon Weather Rev 115:1606–1626
Ropelewski CF, Halpert MS (1989) Precipitation patterns associated with the high index phase of the Southern Oscillation. J Clim 2:268–284
Rusticucci M, Venegas S, Vargas W (2003) Warm and cold events in Argentina and their relationship with South Atlantic and South Pacific Sea Surface temperatures. J Geophys Res Oceans 108(C11):3356. https://doi.org/10.1029/2003JC001793
Rusticucci M, Kyselý J, Almeira G, Lhotka O (2016) Long-term variability of heat waves in Argentina and recurrence probability of the severe 2008 heat wave in Buenos Aires. Theor Appl Climatol 124:679–689. https://doi.org/10.1007/s00704-015-1445-7.
Rusticucci M, Barrucand M, Collazo S (2017) Temperature extremes in the Argentina central region and their monthly relationship with the mean circulation and ENSO phases. Int J Climatol 37:3003–3017. https://doi.org/10.1002/joc.4895.
Rutllant JA (2004) Large-scale atmospheric circulation features associated with the 1997–1999 ENSO cycle and their consequences in the central-Chile precipitation regime. El Niño-La Niña 1997–2000. Their effects in Chile. CONA, Chile, Valparaíso. pp. 61–76.
Rutllant JA, Aceituno P (1991) Southern Hemisphere circulation signals in connection with winter rainfall forecasting in central Chile. International Centre for Theoretical Physics, Trieste, Italy. Internal Report IC/91/64: 20 p.
Saji NH, Yamagata T (2003) Possible impacts of Indian Ocean dipole mode events on global climate. Clim Res 25:151–169. https://doi.org/10.3354/cr025151
Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. Nature 401(23):360–363. https://doi.org/10.1038/43854
Saji NH, Ambrizzi T, Ferraz SET (2005) Indian Ocean dipole mode events and austral surface air temperature anomalies. Dyn Atmos Oceans 39:87–101. https://doi.org/10.1016/j.dynatmoce.2004.10.015
Salio P, Nicolini M, Saulo AC (2002) Chaco low-level jet events characterization during the austral summer season. J Geophys Res 107(D24):4816. https://doi.org/10.1029/2001JD001315
Schwerdtfeger W (1976) Climates of Central and South America. World Surv. Climatol, vol 12. Elsevier Sci, New York, p 522
Seluchi ME, Saulo AC, Nicolini M, Satyamurty P (2003) The northwestern Argentinean low: a study of two typical events. Mon Weather Rev 131:2361–2378. https://doi.org/10.1175/1520-0493(2003)131<2361:TNALAS>2.0.CO;2
Seneviratne SI, Lüthi D, Litschi M, Schär C (2006) Land–atmosphere coupling and climate change in Europe. Nature 443(7108):205–209. https://doi.org/10.1038/nature05095
Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner I, Teuling AJ (2010) Investigating soil moisture–climate interactions in a changing climate: a review. Earth Sci Rev 99(3):125–161. https://doi.org/10.1016/j.earscirev.2010.02.004
Sifuzzaman M, Islam M, Ali MZ (2009) Application of wavelet transform and its advantages compared to Fourier transform. J Phys Sci 13:121–134
Silvestri GE, Vera CS (2003) Antarctic Oscillation signal on precipitation anomalies over southeastern South America. Geophys Res Lett 30(21):2115. https://doi.org/10.1029/2003GL018277
Slingo J, Palmer T (2011) Uncertainty in weather and climate prediction. Phil Trans R Soc A 369:4751–4767. https://doi.org/10.1098/rsta.2011.0161
Solman S, Núñez M (1999) Local estimates of global climate change: a statistical downscaling approach. Int J Climatol 19:835–861
Srivastava A, Sinha Ray K (2000) Prediction of SST anomalies of east Pacific Ocean (Nino 3 region) using a statistical model. Theor Appl Climatol 66:131–138. https://doi.org/10.1007/s007040070020
Statheropoulos M, Vassiliadis N, Pappa A (1998) Principal component and canonical correlation analysis for examining air pollution and meteorological data. Atmos Environ 32:1087–1095
Steinberg D. (2014) Why data scientists split data into train and test. https://info.salford-systems.com/blog/bid/337783/Why-Data-Scientists-Split-Data-into-Train-and-Test. Accessed 27 September 2017.
Tong CHM, Yim SHL, Rothenberg D, Wang C, Lin CY, Chen YD, Lau NC (2018) Assessing the impacts of seasonal and vertical atmospheric conditions on air quality over the Pearl River Delta region. Atmos Environ 180:69–78. https://doi.org/10.1016/j.atmosenv.2018.02.039
Torrence C, Compo GP (1998) A practical guide to wavelet analysis. Bull Amer Meteor Soc 79:61–78
Trenberth KE, Caron JM (2000) The Southern Oscillation revisited: sea level pressures, surface temperatures and precipitation. J Clim 13:4358–4365. https://doi.org/10.1175/1520-0442(2000)013<4358:TSORSL>2.0.CO;2
Vargas W, Penalba O, Minetti J (1999) Monthly precipitation in areas of Argentina and the ENSO: a focus on decision problems. Meteorológica 24:3–22
Vera C, Silvestri G, Barros V, Carril A (2004) Differences in El Niño response over the Southern Hemisphere. J Clim 17:1741–1753. https://doi.org/10.1175/1520-0442(2004)017<1741:DIENRO>2.0.CO;2
Vincent LA, Peterson TC, Barros VR, Marino MB, Rusticucci M, Carrasco G, Ramirez E, Alves LM, Ambrizzi T, Berlato MA, Grimm AM, Marengo JA, Molion L, Moncunill DF, Rebello E, Anunciação YMT, Quintana J, Santos JL, Baez J, Coronel G, Garcia J, Trebejo I, Bidegain M, Haylock MR, Karoly D (2005) Observed trends in indices of daily temperature extremes in South America 1960–2000. J Clim 18:5011–5023. https://doi.org/10.1175/JCLI3589.1
von Storch H, Zorita E, Cubasch U (1993) Downscaling of climate change estimate to regional scales: application to winter rainfall on the Iberian Peninsula. J Clim 6:1161–1171
Wainer I, Prado LF, Khodri M, Otto-Bliesner B (2014) Reconstruction of the South Atlantic subtropical dipole index for the past 12000 years from surface temperature proxy. Sci Rep 4:5291. https://doi.org/10.1038/srep05291
Wei L, Feng Q, Deo RC (2018) Changes in climatic elements in the Pan-Hexi region during 1960–2014 and responses to global climatic changes. Theor Appl Climatol 133(1–2):405–420. https://doi.org/10.1007/s00704-017-2194-6
Whan K, Zscheischler J, Orth R, Shongwe M, Rahimi M, Asare EO, Seneviratne SI (2015) Impact of soil moisture on extreme maximum temperatures in Europe. Weather Clim Extremes 9:57–67. https://doi.org/10.1016/j.wace.2015.05.001

Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics 1:80–83

Yu ZP, Chu PS, Schroeder T (1997) Predictive skills of seasonal to annual rainfall variations in the U.S. affiliated Pacific Islands: canonical correlation analysis and multivariate principal component regression approaches. J Clim 10:2586–2599. https://doi.org/10.1175/1520-0442(1997)010%3C2586:PSOSTA%3E2.0.CO;2

Zamboni L, Mechoso CR, Kucharski F (2010) Relationships between upper-level circulation over South America and rainfall over south-eastern South America: a physical base for seasonal predictions. J Clim 23(12):3300–3315. https://doi.org/10.1175/2009jcli3129.1

Zar JH (1984) Biostatistical analysis, 2nd edn. Prentice-Hall, Inc., Englewood Cliffs, 718 p

Zhang Y, Wallace J, Battisti D (1997) ENSO-like interdecadal variability: 1900–93. J Clim 10:1004–1020. https://doi.org/10.1175/1520-0442(1997)010<1004:ELIV>2.0.CO;2

Zhang W, Lou IC, Kong Y, Ung WK, Mok M (2013) Eutrophication analysis and principal component regression for two subtropical storage reservoirs in Macau. Desalin Water Treat 51:7331–7340. https://doi.org/10.1080/19443994.2013.793921

Zhou H, Liu Y (2017) Spatio-temporal pattern of meteorological droughts and its possible linkage with climate variability. Int J Climatol 38(4):2082–2096. https://doi.org/10.1002/joc.5319

Zitto ME, Barrucand MG, Piotrkowski R, Canziani PO (2016) 110 years of temperature observations at Orcadas Antarctic Station: multidecadal variability. Int J Climatol 36(2):809–823. https://doi.org/10.1002/joc.4384