Article

Statistical Seasonal Rainfall Forecast in the Neuquén River Basin (Comahue Region, Argentina)

Marcela Hebe González [1,2]

[1] Research Center of Sea and Atmosphere (CIMA), CONICET/UBA, UMI-IFAECI/CNRS,
Buenos Aires C1428EGA, Argentina; E-Mail: gonzalez@cima.fcen.uba.ar;
Tel.: +54-11-478-726-93 (ext. 302); Fax: +54-11-478-835-72

[2] Department of Atmospheric Sciences, FCEN, University of Buenos Aires, Intendente Guiraldes 2160,
Buenos Aires C1428EGA, Argentina

Abstract: A detailed statistical analysis was performed at the Neuquén river basin using precipitation data for 1980–2007. The hydrological year begins in March with a maximum in June associated with rainfall and another relative maximum in October derived from snow-break. General features of the rainy season and the excess or deficits thereof are analyzed using standardized precipitation index (SPI) for a six-month period in the basin. The SPI has a significant cycle of 14.3 years; the most severe excess (SPI greater than 2) has a return period of 25 years, while the most severe droughts (SPI less than −2) have a return period of 10 years. The SPI corresponding to the rainy season (April–September) (SPI9) has no significant trend and is used to classify wet/dry years. In order to establish the previous circulation patterns associated with interannual SPI9 variability, the composite fields of wet and dry years are compared. There is a tendency for wet (dry) periods to take place during El Niño (La Niña) years, when there are positive anomalies of precipitable water over the basin, when the zonal flow over the Pacific Ocean is weakened (intensified) and/or when there are negative pressure anomalies in the southern part of the country and Antarctic sea. Some prediction schemes using multiple linear regressions were performed. One of the models derived using the forward stepwise method explained 42% of the SPI9 variance and retained two predictors related to circulation over the Pacific Ocean: one of them shows the relevance of the intensity of zonal flow in mid-latitudes, and the other is because of the influence of low pressure near the Neuquén River basin. The cross-validation used to prove model efficiency showed a correlation of 0.41 between observed and estimated SPI9; there was a probability of detection of wet (dry) years of 80% (65%) and a false alarm relation of 25% in both cases.

1. Introduction

The Comahue region is located in the Central Andes Mountains in Argentina, between 36 °S and 42 °S. The region is crossed by three rivers: the Limay, Neuquén and Negro. With regard to energy, it is one of the major river systems because the hydro-electric energy used in the country is mainly generated there, by operating dams such as El Chocón, Alicurá and Piedra del Aguila in the Limay river basin and Portezuelo, Los Barreales and El Chañar in the Neuquén river basin. It drains an area of 140,000 km^2 and covers almost the whole of the province of Neuquén and part of the provinces of Buenos Aires and Rio Negro. The various uses of this water include hydroelectric generation for the national grid and water supply for the development of local subsistence economies. The operation of dams is highly dependent on rainfall, as it is the main variable that regulates the flow of rivers. In this area the annual cycle of rainfall is characterized by a maximum in winter (April–September) but the interannual rainfall variability is high. Therefore, in order to efficiently operate the dams, it is relevant to forecast the seasonal rainfall. For example, in the Limay and Neuquén river basins, rainfall above average has been recorded in only 4 out of the 13 years during the period 2000–2012. These values are indicative that the region is going through a period of below-normal rainfall which affects the optimal operation of dams. In this context, it becomes relevant to study the relationship between precipitation and flow, as well as the ability to predict these events.

Current forecasts of seasonal rainfall using dynamic modeling have serious shortcomings because of, among other things, the low spatial resolution. That is why the development of statistical rainfall forecasting techniques, specially adapted to the study region, is addressed in this paper. The scientific basis of seasonal climate predictability lies in the fact that slow variations in the earth boundary conditions (*i.e.*, sea surface temperature or soil wetness) can influence global atmospheric circulation, and thus precipitation. Some authors have analyzed these relationships in the southern hemisphere, for example Gissila *et al.* [1] in Ethiopia; Reason [2] in South Africa and Zheng and Frederiksen [3] in Australia. In Argentina, Gonzalez and Vera [4] detected rainfall patterns by analyzing interannual rainfall variability in the Comahue region using a principal component analysis. Gonzalez *et al.* [5] derived rainfall prediction schemes, using multiple linear regressions which explained 51% of the winter rainfall variance in the Limay River basin and 44% in the Neuquén River basin. Gonzalez and Herrera [6] discussed several different statistical models: an autoregressive integrated moving average model (ARIMA), Holt Winter (HW), the Climate Prediction Tool (CPT) and a combination of all three to predict rainfall in southern Argentina (Patagonia). Scarpati *et al.* [7] and Gonzalez [8] studied rainfall and flow characteristics in this region of Argentina. Gonzalez and Dominguez [9] applied a prediction scheme using multiple linear regression and showed that 46% of the standardized precipitation index (SPI) variance could be explained by this model in the Limay river basin. In general, the SPI was used to quantify the conditions of deficit or excess of precipitation for a several-month interval, and it is usually used in hydrological studies (Mckee *et al.* [10,11]), especially for improving the operation of dams. The aim of this paper is to develop a technique to predict seasonal rainfall in the Neuquén River basin using statistical methods applied to

SPI. This basin is located in the northern part of the Comahue region. The Neuquén River is the boundary of the Neuquén and Mendoza provinces. The river runs from the high Los Andes Mountain in the west towards the Patagonian plateau in the east and drains an area of 30,000 km^2. Between 80%–90% of the total rainfall occurs in winter, and this is the most important factor that contributes to river flow. Snowfall accumulates until late spring in high mountains in the west and generates a secondary peak in river flow. The paper is organized as follows: Section 2 describes the dataset and the methodology. Section 3.1 presents the SPI low frequency features in Neuquén river basin. Section 3.2 shows the difference in the behavior of sea surface temperature and atmospheric circulation when extreme SPI is registered. Section 3.3 presents some regression models to estimate wet and dry events and discusses their efficiency. Section 4 presents the main conclusions.

2. Data and Methodology

Rainfall data from different sources (National Meteorological Service, the Secretary of Hydrology of Argentina and the Territory Authority of the Limay, Neuquén and Negro river basins) in 12 stations of the Neuquén River basin (NRB) over the period 1980–2007 were used in this study (Figure 1). In this period, all the stations have less than 20% missing monthly rainfall data and their quality has been carefully ascertained. A number of techniques were applied as part of the quality control process. First we discriminated cases with no precipitation in one month due to missing data. No stations have records affected by changes of location and instrumentation. Rainfall greater than the 95th percentile was evaluated in order to detect outliers. In a consistency check, the observation was compared with a nearby station value to see if it was physically or climatologically consistent. Suspicious observations due to inconsistencies were removed. Precipitation was homogeneous all over the basin as proven by applying the Lund method using a 0.6 correlation coefficient (Garbarini and González [12]). In order for the average of monthly precipitation to be representative of the precipitation over the basin, a mean rainfall series was calculated in NRB.
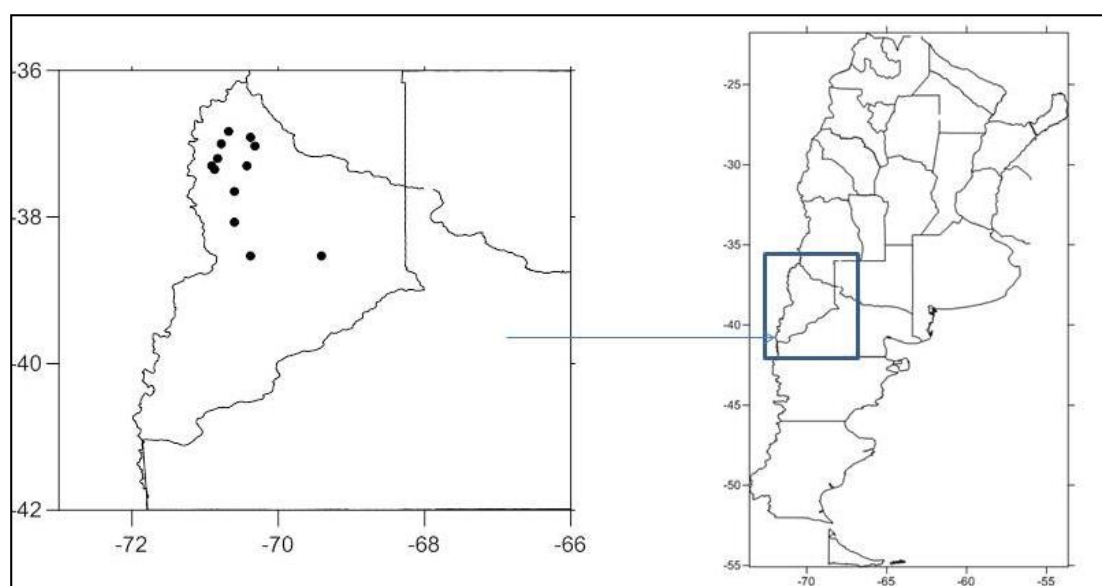

Figure 1. Area of study and location of the stations used.

The SPI was defined by Mckee [10] with the aim of detecting wet and dry periods. The basic approach is to use standardized precipitation for a set of time scales which together represent water sources of several types. The accumulated precipitation in different time scales is fitted using a gamma distribution. After fitting the gamma probability distribution to precipitation data, the cumulative distribution is transformed through an equal-probability transformation into a normal distribution so that the mean SPI is set to zero. It represents the number of standard deviations that the precipitation is away from its historical average. Negative (positive) values of SPI represent deficits (excess) of rainfall compared with the expected normal value. The period used to calculate SPI depends on the use. For example, SPI calculated at a 6-month scale is useful for evaluating dry and wet periods and is useful for hydrological droughts. The one-month SPI has been useful for evaluating agronomical droughts (Mckee [10]). The main advantages of using SPI are the simplicity of calculation because it only uses rainfall data, the possibility of calculating the index at different time scales, and the consistent frequency of extreme and severe cases for any location and any timescale because of the normal distribution. Mckee [10] suggests a classification of excess and dry cases using SPI (see Table 1). Hayes *et al.* [13] explained the advantages and limitations of SPI use and Heim [14] did a detailed comparison of different drought indexes used including SPI.

Table 1. Years classified using the value of SPI9. Percentage frequency is detailed.

	Category	Definition	Frequency (%)	Years
DRY CASES	Extreme drought	$SPI9 < -2$	3.5	1998
	Severe drought	$-2 < SPI9 < -1.5$	3.5	1996
	Moderate drought	$-1.5 < SPI9 < -1$	10.7	1989, 1990, 2007
	Slight drought	$-1 < SPI9 < -0.5$	7.1	1985, 2003
NORMAL CASES	Normal	$-0.5 < SPI9 < 0.5$	39.3	1981, 1983, 1984, 1987, 1988, 1991, 1992, 1994, 1995, 1999, 2004
WET CASES	Slight excess	$0.5 < SPI9 < 1$	21.4	1986, 1993, 1997, 2000, 2002, 2005
	Moderate excess	$1 < SPI9 < 1.5$	10.7	1980, 2001, 2006
	Severe excess	$1.5 < SPI9 < 2$	3.5	1982
	Extreme excess	$SPI9 > 2$	0	-

In this work, SPI was calculated using monthly rainfall data at a time scale of six months for the period 1980–2007. That means that SPI was calculated for every month using the accumulated precipitation in the six-month period previous to that month. A gamma probability density function is fitted to calculate SPI. SPI greater than zero indicates water excess, while a value lower than zero indicates a deficit. The magnitude of the index allows the six-month accumulated rainfall to be classified into categories that go from extreme drought to extreme excess (extremely dry SPI less than −2; severe dry, SPI between −2 and −1.5; moderate dry, SPI between −1.5 and −1; slight dry, SPI between −1 and −0.5; normal, SPI between −0.5 and 0.5; slight wet, SPI between 0.5 and 1; moderate wet, SPI between 1 and 1.5; severe wet, SPI between 1.5 and 2 and extremely wet, SPI greater than 2).

Some characteristics of the SPI series were detailed with the aim of describing its behavior: the possible long term trend, the cycles greater than a year and the probability that an extreme value occurred. All these factors are important when the interannual variability is studied because they influence the amount of precipitation in a particular year and they must all be considered together.

SPI series was analyzed for low frequency variability using a linear trend method of minimum squares, and statistical significance was tested using a Student T test. The dominant periodicities associated with periods greater than one year were identified through spectral analysis. The technique used was the Blackman-Tukey method (Mitchell *et al.* [15], Blackman and Tukey [16]), which consists of applying a harmonic analysis to the autocorrelation function in order to detect cycles present in the original series. The spectral analysis was applied to the series after filtering the annual cycle. The aim is to find the cycles with the greatest spectral densities, that is, the frequency regions, consisting of many adjacent frequencies that mostly contribute to the overall periodic behaviour of the series. Therefore, a smoothing of spectral values using a weighed moving average transformation was done using a Tukey window. Its significance, depending on whether the spectrum fits a white or a red noise model, was also tested using the Blackman-Tukey method, and a confidence level of 95% was used to detect significant peaks. According to Blackman and Tuckey [16] each spectral estimate is distributed as a chi-square divided by degree of freedom. This fact allows the confidence interval of each estimate of a computed spectrum to be easily determined by reference to a chi-square table.

Sometimes return periods are used to express the expected frequency of a flood or droughts of a certain magnitude in order to evaluate the actual risk in a basin. To calculate the return period of extreme SPI, a Gumbel function was fitted to the maximum and minimum monthly SPI.

As SPI was calculated for every month of the six-month period from 1980 to 2007, the value of SPI in September (SPI9) represents the accumulated rainfall from April to September. SPI9 is used in this study because rainfall season takes place over this period in the Neuquén River basin. Years were classified as wet or dry—as well as their different categories—using the value of SPI9, and represent the accumulated rainfall from April to September. Composite fields in April (when the rainfall period begins) of several variables for wet and dry years and the difference between them were plotted. Variables used were monthly sea surface temperatures (SST), 500 hPa (G500), 1000 hPa (G1000) and 200 hPa (G200) geopotential heights, zonal (U) and meridional (V) winds at 850 hPa and precipitable water (PW) from National Center of Environmental Prediction (NCEP) reanalysis (Kalnay *et al.* [17]). Monthly anomalies were determined by removing the climatologically monthly means from the original values.

SPI9 predictors were defined using the difference between meteorological variables behavior in wet and dry years. Different sets of predictors, carefully selected based on statistical significance and physical reasoning, were proved to design statistical forecast models using the forward stepwise regression method (Wilks [18]) which retained only the variables correlated with a 95% confidence level. Forward stepwise regression is a model-building technique that finds subsets of predictor variables that most adequately predict responses on a dependent variable by linear regression, considering the specified criteria for adequacy of model fit (Darlington [19]). A common approach to better estimate predicting skills is a cross-validation (Wilks [18]), where n − 1 years were used for calibration and the remaining year was used to validate the model. This process was repeated several times with different years. This method is generally strong in the presence of long-term climate variability and it is used especially when the amount of data is not so large. In the process of cross-validation, the exclusion of a single year from

the data might not cause important changes in the model when detecting any evidence of numerical instability. With the aim of evaluating forecast effectivness, two indices (Wilks [18]) were also calculated for wet and dry cases. The probability of detection (POD) is defined as the fraction of those occasions when the forecast event occurred in which it was also forecast. The false alarm relation (FAR) is the proportion of forecast events that fail to happen.

3. Results and Discussion

3.1. Low Frequency Standarized Precipitation Index Features in the Neuquén River Basin

In this section some characteristics of SPI series are described in order to better understand the behavior of this variable at different time scales such as the evolution in log periods (trend), the possible periodicities included in time series, and the probability that an extreme SPI occurs. SPI was calculated for every month of the six-month period from 1980 to 2007. For each year, the maximum and the minimum SPI were considered and the return periods were calculated. Figure 2a shows the maximum SPI *vs.* the return period (in years). It can be seen that extreme water excess, related to floods (classified by Mckee [10] as SPI greater than 2) has a return period of almost 25 years. Meanwhile, extreme minimum SPI (less than −2), associated with droughts, has a return period of only 10 years (Figure 2b). Therefore, there is higher probability of droughts than floods in the NRB.
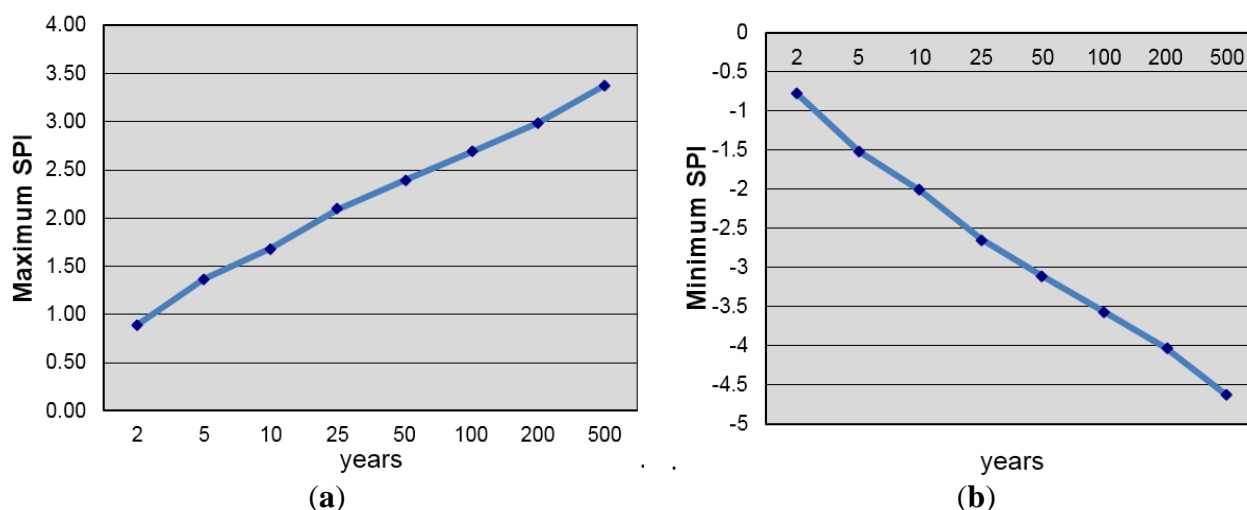

Figure 2. Maximum (**a**) and minimum (**b**) six-month standardized precipitation index in the Neuquén River basin *vs.* return period (in years).

A six-month SPI was considered to perform a spectrum analysis in order to detect significant periodicities greater than one year. Only a significant peak of about 14.3 years was detected (Figure 3).

Figure 4 shows the mean (1980–2007) monthly rainfall in NRB. It shows that precipitation has an important annual cycle with a maximum in winter. As the most significant rainfall amount occurs from April to September, the analysis of SPI9 (a six-month SPI for the precipitation accumulated over this period) will be detailed hereinafter. Linear trend approximation was used to evaluate the observed change in SPI9 series (Figure 5) and a negative but non-significant (95% confidence level) linear trend was detected.
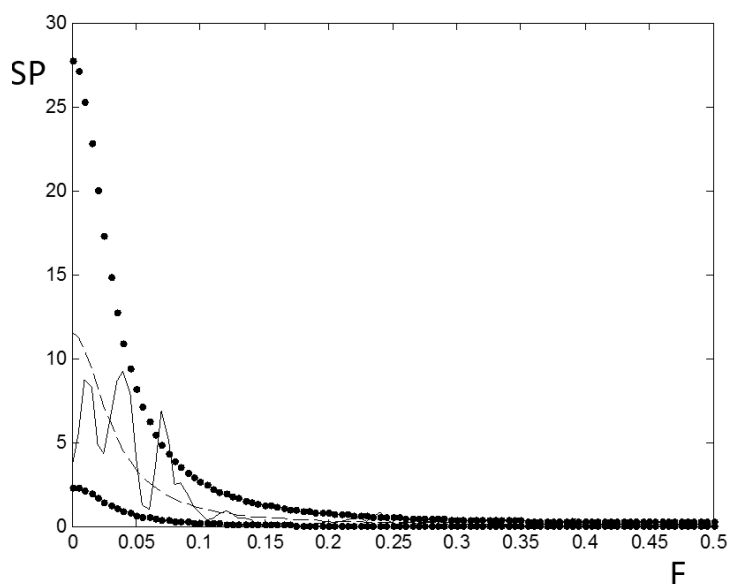
Figure 3. Spectral analysis of the six-month standardized precipitation index in the Neuquén River basin. SP is the spectral density and F is the frequency (year^{-1}).
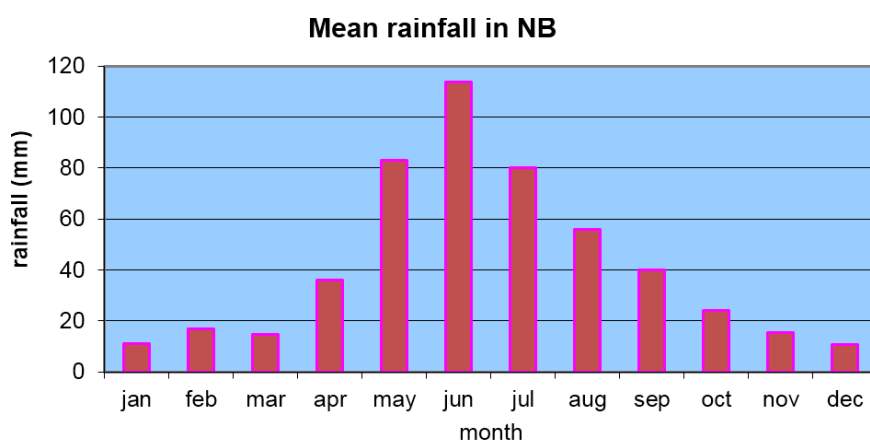

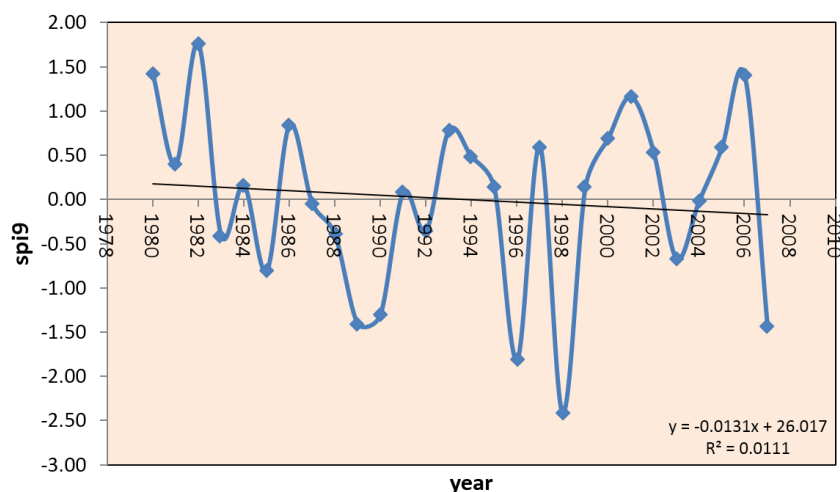**Figure 4.** Mean annual rainfall (1980–2007) in the Neuquén River basin.

Figure 5. Evolution of six-month-standardized precipitation index in September in the Neuquén river basin. Linear trend in black.

3.2. Atmospheric Circulation and Sea Surface Temperature Patterns Associated with Extreme Rainfall in the Neuquén River Basin

As the most important quantity rainfall amount occurs from April to September (Figure 4), the value of SPI9 (SPI for the precipitation accumulated in this period) will be used to classify years into different categories: dry, normal and wet cases (Table 1).

Anomaly composite fields for wet and dry cases in April and the difference between dry and wet composites are calculated for several meteorological variables in order to detect possible previous patterns associated with the occurrence of extreme events. Figure 6 [17] shows G1000 composites for dry (Figure 6a) and wet (Figure 6b) cases and the difference between them (Figure 6c). It can be seen that subtropical high and subpolar low are especially intensified in eastern Pacific Ocean in dry years. Therefore, the westerlies are intensified and there is less energy exchanged between low and middle latitudes, so the rainfall systems move eastward from the Pacific Ocean with little movement towards the north over the area of study. The main feature, however, is the east-west dipole that can be observed in Figure 6c, related to the presence of G1000 negative (positive) anomalies southern Argentina in the Antarctic Sea, centered in 70°S and 60°W in wet (dry) cases. Negative G1000 anomalies reinforce frontal rainfall systems that move from the Pacific Ocean, thus enhancing rainfall. This pattern is also present in middle (G500) and high (G200) levels (figures not shown).

Figure 7 [17] shows the difference in U (Figure 7a), V (Figure 7b) and PW (Figure 7c) between dry and wet composites. Figure 7a reflects the fact that westerlies are especially intensified in dry cases (see positive U anomalies centered in 47°S and 150°W). This agrees with the pattern described for G1000. Figure 7b shows a northerly flow over NRB, intensified in wet years compared with dry ones. Figure 7c shows that the moisture availability over the basin is greater in wet than in dry years. This fact contributes to more intense moist air advection from the north in wet years.

3.3. Statistical Models to Predict Extreme Rainfall in the Neuquén River Basin.

Predictors that represent the statistical associations between SPI9 and circulation or SST patterns are defined using the difference composite fields described in the previous section. It is important to point out that they have been carefully selected, with regard to their physical reasoning. Taking into account the results detailed in the previous section, the variables used as predictors were sea surface temperature and geopotential height over the tropical Pacific Ocean because they are related to the wave train from the west associated with rainfall systems in the area of study. In fact, a warming or cooling of some region of the oceans can act as a remote forcing and generate teleconnections. Indeed, the most relevant SST pattern in the Pacific Ocean is the El Niño-Southern Oscillation. The SST anomalies in tropical Pacific generate a Rossby wave trend which propagates meridionally towards middle-latitudes from the tropical source (Kidson [20], Mo [21], Nogues Paegle and Mo [22]). This pattern, called the "Pacific South American Pattern", is described by Mo [21] when depicting the South Hemisphere climate. Other variables used as predictors were sea surface temperature over the Pacific and Atlantic Oceans near the continent, precipitable water and regional wind near the area of study. The latter were chosen because they are related to moist air advection that can increase rainfall amounts.
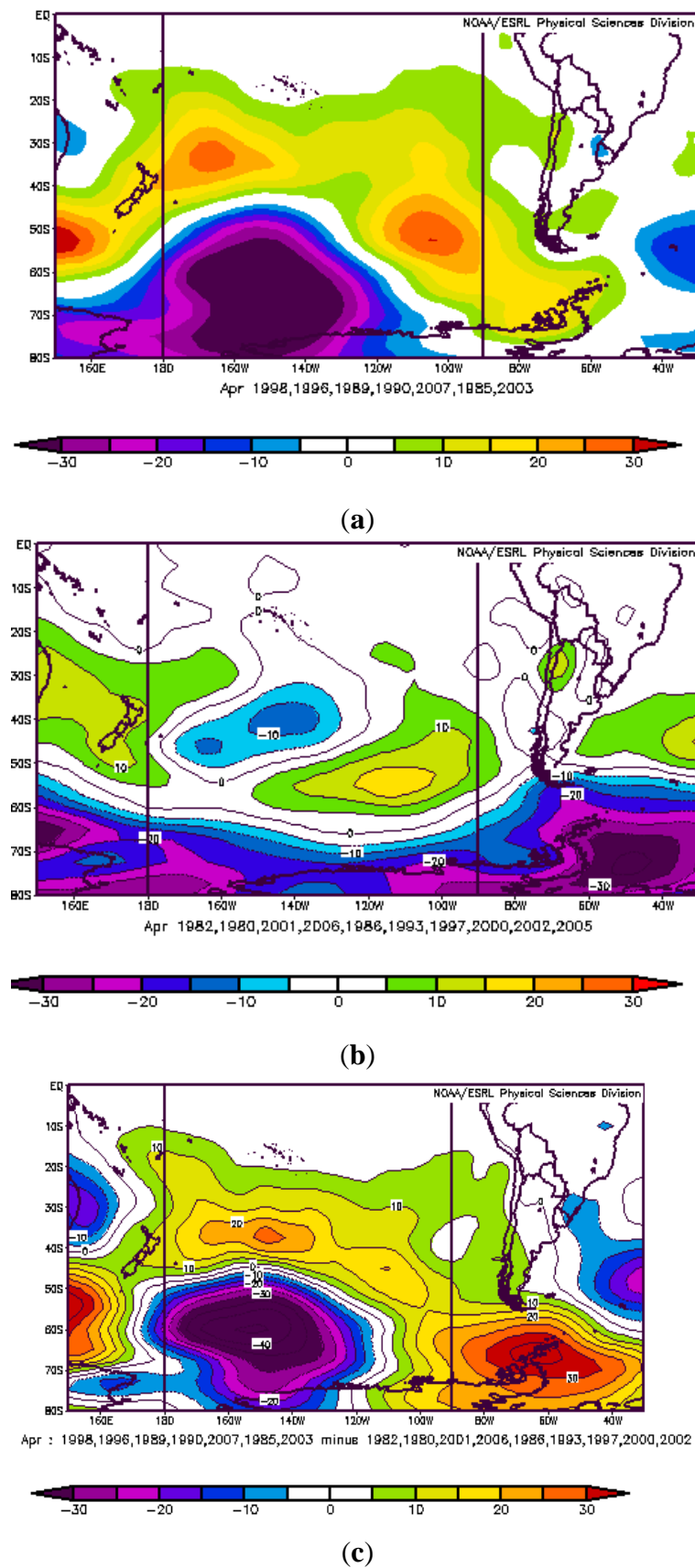
Figure 6. 1000 hPa geopotential height (in m) composites for (**a**) dry and (**b**) wet cases and (**c**) the difference between them following the classification showed in Table 1.
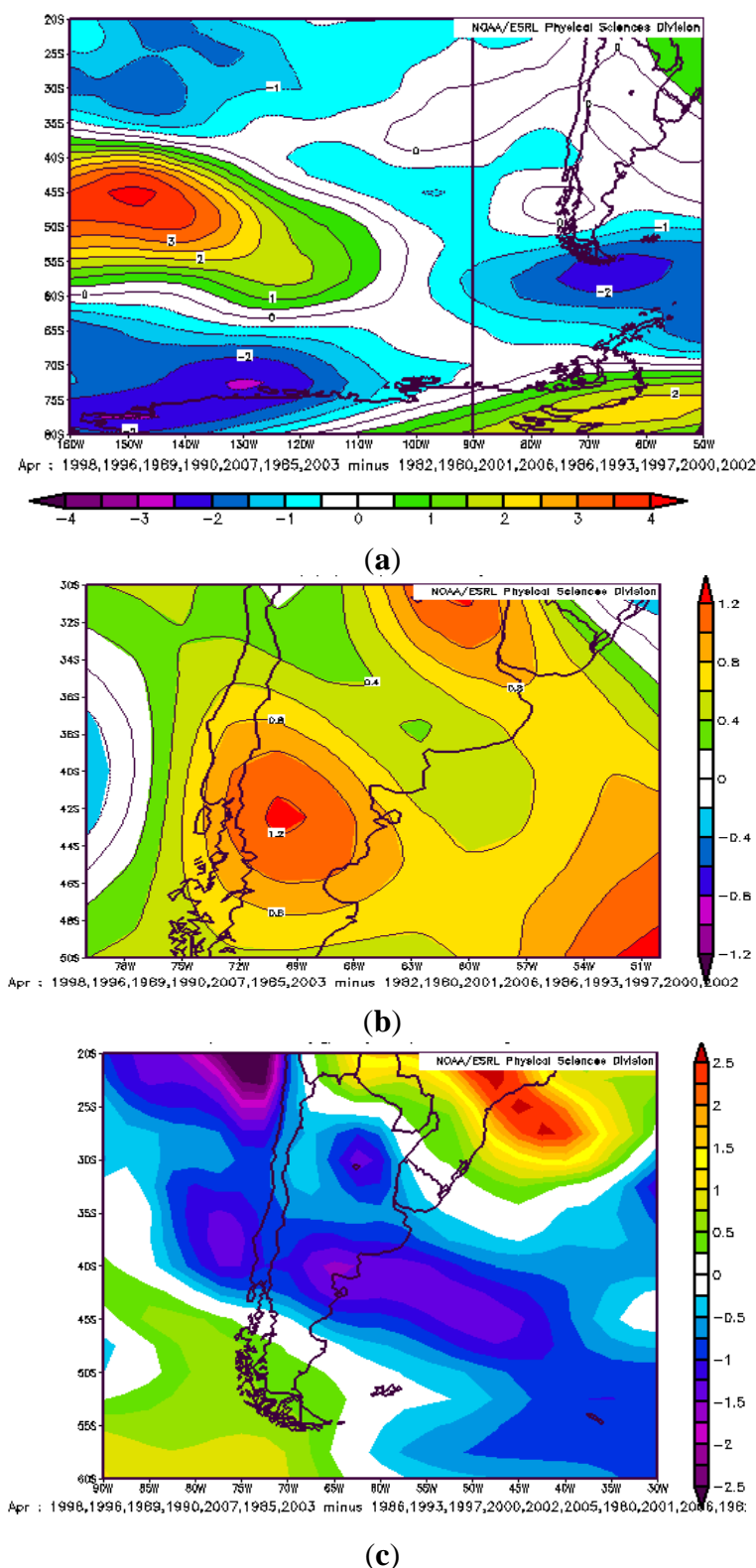
(**a**)

(**b**)

(**c**)

Figure 7. (**a**) 850 hPa zonal wind (in m/s), (**b**) 850 hPa meridional wind (in m/S) and (**c**) precipitable water (in Kg/m^2) difference composites between dry and wet years following the classification showed in Table 1.

Figure 8 shows an SST composite for dry (Figure 8a [17]) and wet (Figure 8b [17]) years and the difference between them (Figure 8c). The difference is greater in two major regions associated with

El Ni ño-Southern Oscillation (ENSO): the tropical Pacific Ocean, especially in the eastern Pacific and southeastern Pacific near the cost of South America. Wet years are related to the warm phase of ENSO (positive SST anomalies in the eastern tropical Pacific and negative anomalies in the southeastern Pacific Ocean), while dry years tend to occur in the cold ENSO phase.
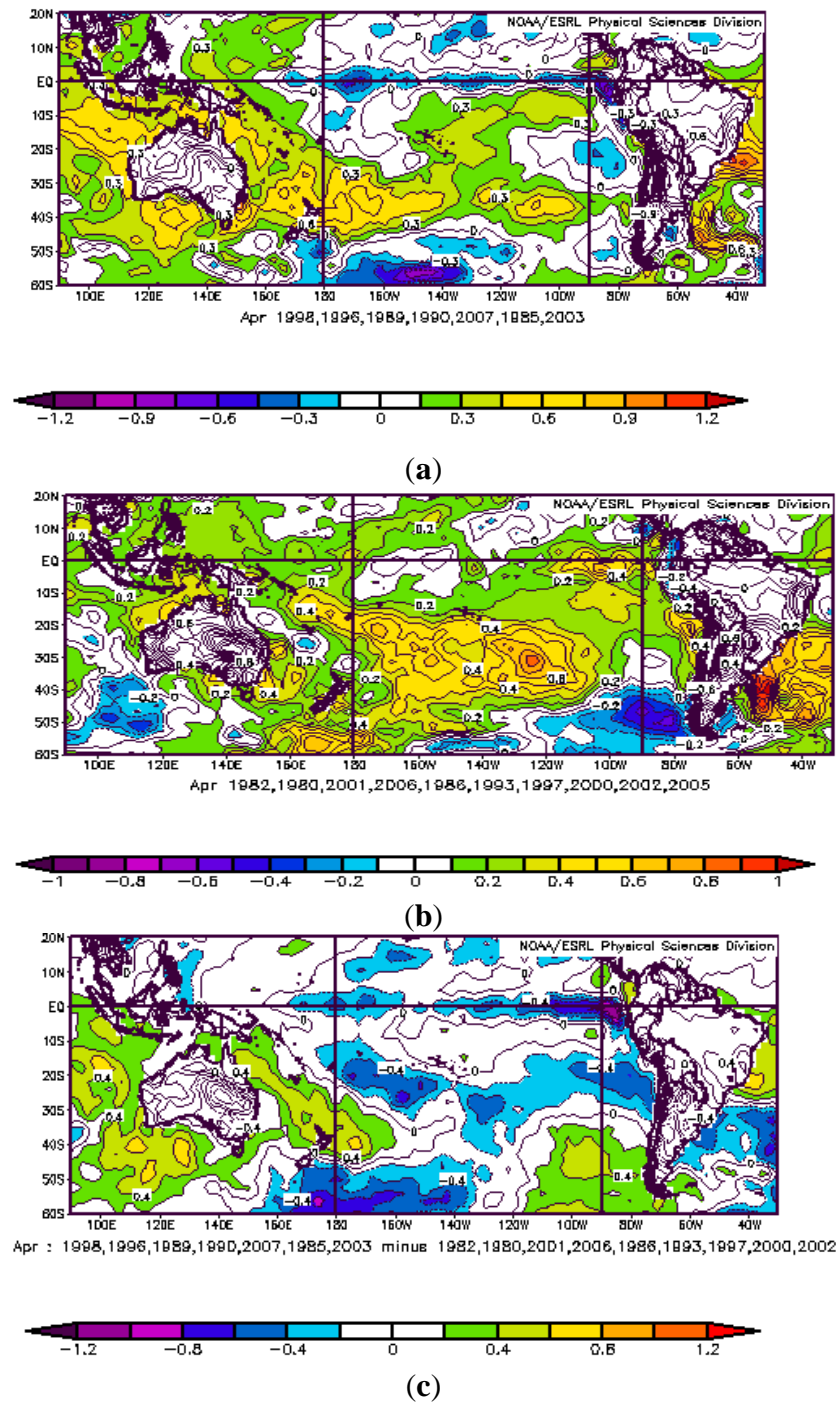

Figure 8. Sea surface temperature (in °C) composites for (**a**) dry and (**b**) wet cases and (**c**) the difference between them following the classification showed in Table 1.

Predictors are defined as their mean values in areas where significant differences between wet and dry cases in April were addressed. Their definitions are detailed in Table 2. Note that the two predictors are also defined which reflect the north-south (DNS) and east-west (DEW) dipoles showed in the G1000 difference composites field (Figure 6c). DNS represents the difference between subpolar low and subtropical high anomalies and DEW represents the difference in geopotential height anomalies in southern Argentina for dry and wet years, both shown in Figure 6c. These predictors are defined in G1000 representing low levels (subscript "l" in Table 2), in G500 representing middle levels ("m" subscript) and in G200 representing high levels ("h" subscript). The correlation between the defined predictors and SPI9 is calculated and values greater than 0.37 are statistically significant (95% confidence level).

Table 2. Predictors defined using difference between dry and wet composites fields (*). Correlations greater than 0.37 are significant at 95% confidence level (**bold**).

Predictor	Variable	Area	Correlation
P1l	G1000	(170 °, 140 °O; 55 °, 70 °S)	**0.39**
P1m	G500	(170 °, 140 °O; 55 °, 70 °S)	**0.45**
P1h	G200	(170 °, 140 °O; 55 °, 70 °S)	0.35
P2l	G1000	(150 °. 140 °O; 35 °, 40 °S)	**−0.49**
P2m	G500	(150 °, 140 °O; 35 °, 40 °S)	**−0.46**
P2h	G200	(150 °, 140 °O; 35 °, 40 °S)	**−0.4**
P3l	G1000	(80 °, 60 °O; 60 °, 75 °S)	−0.33
P3m	G500	(80 °, 60 °O; 60 °, 75 °S)	**−0.47**
P3h	G200	(80 °, 60 °O; 60 °, 75 °S)	**−0.43**
DNSl	G1000	P1l–P2l	**0.55**
DNSm	G500	P1m–P2m	**0.54**
DNSh	G200	P1h–P2h	**0.43**
DEWl	G1000	P1l–P3l	**0.47**
DEWm	G500	P1m–P3m	**0.57**
DEWh	G200	P1h–P3h	**0.51**
U1	U	(155 °, 140 °O; 45 °, 50 °S)	**−0.46**
V1	V	(72 °, 68 °O; 42 °, 44 °S)	−0.13
PW1	PW	(65 °, 55 °O; 40 °, 45 °S)	0.32
SST1	SST	(90 °, 80 °O; 5 °N, 5 °S)	−0.15
SST2	SST	(100 °, 90 °O; 40 °, 50 °S)	**−0.37**
DEWm	G500	P1m–P3m	**0.57**

(*) G is geopotential height. G, DNS and DEW are defined in low (G1000), middle (G500) and high (G200) levels.

Different sets of predictors were carefully selected, based on statistical significance and considering that all predictors must be independent from each other. They were used to design statistical forecast models using the forward stepwise regression method for predicting SPI9. A cross-validation method was applied and the stability and efficiency of all the models were proven. Table 3 shows the different sets of predictors used to run the linear regression model using the forward stepwise technique, the variance of SPI9 explained by the model, the predictors that the technique selected and the correlation between observed SPI9 and SPI9 derived from cross-validation. The first model explained only 32% of

the SPI9 variance but the others explained between 42% and 43%. Table 4 details the linear regression equation derived from each set of predictors. The first model selected the DEW predictor which indicates the importance of the geopotential height anomalies in southern Argentina. The second mainly showed the influence of the difference between subpolar low and subtropical high anomalies, as predictors P2 and P3 were selected. The third and fourth model include both effects as DNS and P3 or DEW and P2 predictors were selected respectively. It is important to notice that the predictors selected by the forward stepwise technique were all representative of circulation patterns (geopotential heights and combinations) and the predictors defined with SST and PW were not selected by the method. This indicates the great importance of previous circulation patterns in the production of rainfall over the NRB. The correlations between observed SPI9 and the SPI derived from cross-validation method are significant at 95% confidence level for all the models.

Table 3. Sets of predictors used in linear regression model.

Set	Predictors Entered in the Model	Explained Variance	Selected Predictors	Correlation between Observed and Predicted Spi9
Set 1	DNSm DEWm PW SST2	32	DEWm	0.42
Set 2	P1m P2m P3m SST2 PW	43	P2m P3m	0.36
Set 3	DNSm P3m PW SST2	42	DNSm P3m	0.42
Set 4	DEWm P2m PW SST2	42	DEWm P2m	0.44

Table 4. Linear regression equations derived from each set of predictors using the forward stepwise method.

Set	Regression Equation	F	*p*
Set1	0.007 DEWm + 0.001644	12.435	0.001
Set 2	−0.0132 P2m − 0.0101 P3m + 0.001644	9.75	0.0007
Set 3	0.0057 DNS − 0.008 P3m + 0.001644	9.07	0.001
Set4	−0.009 P2m + 0.006 DEWm + 0.001644	9.33	0.0009

A measure of strength of the regression is the F-ratio, defined as the relationship between the mean square regression and the mean square error (Wilks [18]). It is worthwhile that the F-ratio was high because a strong relationship between SPI9 and the predictors will produce a large mean square regression and small mean square error. As the residuals of the regression are independent and follow a normal distribution, under the null hypothesis of no lineal regression, the F-ratios detailed in Table 4 for each model indicate that the regression models provide reasonably good forecast with 95% of confidence. Figure 9 shows the observed SPI9 and the SPI derived from cross-validation method applied to each one of the models detailed in Table 4. The mean (ensemble) of all models is also detailed in the figure. In general, the predicted SPI9 tends to underestimate the extreme observed values.

The effectiveness of the different models was measured in semi-quantitative terms. The observed and forecast SPI9 were categorized in two categories, SPI9 greater than zero ("above" category) and SPI9 lower than zero ("below" category), and they were compared by constructing a contingency table. These contingency tables differ significantly from a random one, using a Chi-square test in all cases. Two measures of accuracy were calculated for both categories and for each one of the models: POD and FAR (Table 5). The probability of detection (POD) is defined as the fraction of those occasions when the

forecasted event occurred on which it was also forecast. The false alarm relation (FAR) is the proportion of forecast events that fail to happen. All POD values were greater than 60% but POD was better for "above" cases than for "below". All FAR values were less than 38% but FAR values were better for "below" cases rather than "above" cases. Models derived from set 1 and set 4 are better than those from sets 2 and 3. If we use the ensemble of all models, 86)% of the wet and 71% of the dry cases would be detected, and 17% of the dry and 25% of the wet cases would result in false alarm cases.
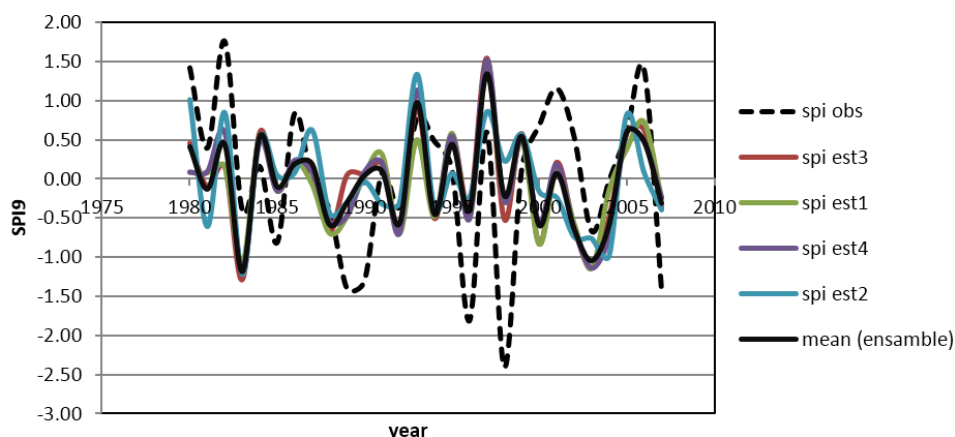

Figure 9. Observed 6-month standardized precipitation index in September and 6 months-standardized precipitation index derived from cross-validation applied to each one of the models detailed in Table 4. The mean (ensemble) of all them is also drawn.

Table 5. POD and FAR calculated for the categories "above" (SPI9 > 0) and "below" (SPI9 < 0).

POD	Below (SPI < 0)	Above (SPI > 0)	FAR	Below (SPI < 0)	Above (SPI > 0)
set1	0.79	0.93	set1	0.08	0.19
set2	0.6	0.77	set2	0.25	0.38
set3	0.69	0.8	set3	0.25	0.25
set4	0.77	0.87	set4	0.17	0.19
ensamble	0.71	0.86	ensamble	0.17	0.25

4. Conclusions

In this paper, rainfall accumulated during the rainy season (April–September), represented by SPI9, was studied specifically for the Neuquén River basin. The main goal of this work is to detect previous (April) circulation and sea surface temperature patterns which could be used as predictors for a possible extreme rainy season. Several approaches to estimate the quantitative value of SPI9 were developed. Although there is a tendency for wet (dry) periods to take place in El Niño (La Niña) years, the statistical linear models proved that the circulation over the Pacific Ocean and over the basin are the main factors which influence rainfall. The linear models derived resulted in efficient forecasting of SPI9, and the probability of detecting an extreme case was greater than 60%. These results are important to operate dams in the Comahue region because the statistical models may be useful to improve the forecast required. This paper is meant to contribute to greater knowledge of climate variability, in order to make better seasonal predictions for hydrological applications.

Acknowledgments

References

1. Gissila, T.; Black, E.; Grime, D.; Slingo, J.M. Seasonal forecasting of the Ethiopian summer rains. *Int. J. Climatol.* **2004**, *24*, 1345–1358.
2. Reason, C. Subtropical Indian Ocean SST dipole events and southern Africa rainfall. *Geophys. Res. Lett.* **2001**, *28*, 2225–2227.
3. Zheng, X.; Frederiksen, C. A study of predictable patterns for seasonal forecasting of New Zealand rainfall. *J. Clim.* **2006**, *19*, 3320–3333.
4. González, M.H.; Vera, C.S. On the interannual winter rainfall variability in southern Andes. *Int. J. Climatol.* **2010**, *30*, 643–657.
5. González, M.H.; Skansi, M.M.; Losano, F. A statistical study of seasonal winter rainfall prediction in the Comahue region (Argentine). *Atmosfera.* **2010**, *23*, 277–294.
6. González, M.H.; Herrera N. Statistical prediction of winter rainfall in Patagonia (Argentina). In *Horizons in Earth Science Research*, 1st ed.; Veress, B., Szigethy, J., Eds.; NOVA Science Publishers: Hauppauge, NY, USA, 2014; Volume 11, pp. 221–238.
7. Scarpati, O.; Kruse, E.; Gonzalez, M.H.; Ismael, A.; Vich, J.; Capriolo, A.; Caffera, R. Updating the Hydrological Knowledge: A case of Study. In *Handbook of Engineering Hydrology, Environmental Hydrology and River Management*, 1st ed.; Eslamian, S., Ed.; Taylor & Francis Group LLC: Boca Raton, FL, USA, 2014; Volume 3, pp. 443–457.
8. González, M.H. Some indicators of interannual rainfall variability in Patagonia (Argentine). In *Clim. Variability—Regional and Thematic Patterns*, 1st ed.; Tarhule, A., Ed.; INTECH: London, UK, 2013; pp. 133–161.
9. González, M.H.; Dominguez, D. Statistical prediction of wet and dry periods in the Comahue Region (Argentina). *Atmos. Clim. Sci.* **2012**, *2*, 2160–2414.
10. Mckee, T.; Doesken, N.; Kleist, J. The relationship of drought frequency and duration to times scales. In Proceedings of the Eighth Conference of Applied Climatology, Anaheim, CA, USA, 17–22 January 1993.
11. McKee, T.; Doesken, N.; Kleist, J. Drought monitoring with multiple time scales. In Proceedings of the Ninth Conference on Applied Climatology, Dallas, TX, USA, 15–20 January 1995.
12. Garbarini, E.M.; Gonzalez, M.H. Variabilidad de la Precipitación en Argentina: Factores de Influencia. In the Proceedings of 2do Congreso Internacional de Hidrología de Llanuras, Santa Fe, Argentina, 23–26 September 2014.
13. Hayes, M.; Svoboda, M.; Wilhite, D.; Vanyarkho, O. Monitoring the 1996 drought using the standardized precipitation index. *Bull. Am. Meteorol. Soc.* **1999**, *80*, 429–438.

14. Heim, R. A review of Twentieth Century drought indices used in the United States. *Bull. Am. Meteorol. Soc.* **2002**, *83*, 1149–1165.

15. Mitchell, J.M.; Dzerdzeevskii, B.; Flohn, H.; Hofmeyr, W.L.; Lamb, H.H.; Rao, K.N.; Wallén, C.C. *Climatic Change*; Technicall Note No. 79; World Meteorological Organization (WMO): Geneva, Switzerland, 1966.

16. Blackman, R.B.; Tukey, J.W. *The Measurement of Power Spectral from the Point of View of Communication Engineering*; Dover: New York, NY, USA, 1958.

17. Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; *et al.* The NCEP/NCAR reanalysis 40 years-project. *Bull. Am. Meteor. Soc.* **1996**, *77*, 437–471.

18. Wilks, D. *Statistical Methods in the Atmospheric Sciences: An Introduction*, 2nd ed.; Elsevier: San Diego, CA, USA, 1995.

19. Darlington, R.B. *Regression and Linear Models*; McGraw-Hill: New York, NY, USA, 1990.

20. Kidson, J. Principal modes of southern hemisphere low frequency variability obtained from NCEP-NCAR reanalyses. *J. Clim.* **1999**, *1*, 1177–1198.

21. Mo, K.C. Relationships between low frequency variability in the southern Hemisphere and sea surface temperature anomalies. *J. Clim.* **2000**, *13*, 3599–3610.

22. Nogues Paegle, J.; Mo, K.C. Linkages between summer rainfall variability over South America and sea surface temperature anomalies. *J. Clim.* **2002**, *15*, 1389–1407.