INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2015)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.4439

RMetS
Royal Meteorological Society

Climate characteristics and their relationship with soybean and maize yields in Argentina, Brazil and the United States

María Paula Llano* and Walter Vargas
National Scientific and Technological Research Council (CONICET), Department of Atmospheric and Oceanic Sciences, Faculty of Sciences, University of Buenos Aires, Argentina

ABSTRACT: The impact of climate variability on exportable surpluses of maize and soya beans in producing/exporting countries is well known. This situation necessitates the study of climate variability and yields in some of the major producer countries (Argentina, Brazil and the United States). This study aims to characterize the climate regimes of each region, seeking the highest degree of regional homogeneity and representativeness of yields and climate variables. The series variability is analysed according to various effects (different stages of crop development and conditioning of yield). The temperatures and rainfall occurring during the flowering stage exert greater conditioning on yield at the end of the crop cycle. Different theoretical fit models of yield are obtained for both crops with climate information from their core areas of high production.

KEY WORDS precipitation; extreme temperature; crops yields

Received 28 May 2014; Revised 12 May 2015; Accepted 10 June 2015

1. Introduction

The climate and the crop yields of maize and soya bean are closely related (Thompson, 1969). The influence of climate on a given crop depends on, among other factors, geographic location and production conditions. Climatic variability is a major source of uncertainty in estimating crop development. It is important to know the impact of climate variability on the exportable surpluses of maize and soya bean in producing and exporting countries, as both crops produce raw materials for food, fuel and industry. This situation forces the study of climate variability and yields in major producing and exporting countries.

Statistics from the Food and Agriculture Organisation (FAO, division of the United Nations) for maize production (1979–2009) indicate that the world largest producer is the United States, averaging more than 200 million tons. The second largest producer is China, with 100 million tons. Third place is held by Brazil, with 30 million tons per year on average over 30 years. Argentina ranks sixth, behind Mexico and France, with the three countries producing close to 15 million tons a year. For the cultivation of soya bean, according to the average production values during this period, the United States is the largest producer, with 64 million tons. Second is Brazil with 30 million, and Argentina is third in the rankings with 17 million tons of soya bean. Based on these statistics and on export capacity, the countries selected for this study are Argentina, Brazil and the United States.

Within applied climatology, there is an important area of agronomic research that studies the connections between climate and the productivity of cereals (Kucharik and Serbin, 2008). Trends and fluctuations in yields are relevant to micro- and macro-economy decision makers, as well as food-security programmes, particularly in developing countries. For this reason, the evaluation of the potential economic benefits of understanding climate variability may favour the formation of decision frameworks in agribusinesses (Sonka *et al.*, 1998).

During the vegetation period of plants, climate conditions (e.g. precipitation and temperature) affect crop growth, causing variations in yields. The described effects of the influence of climate demand a continuous effort to improve agricultural technology and to generate a management strategy and thereby reduce the negative impacts of climate (Hu and Buyanovsky, 2003). Understanding and assessing the influence of climate on crop yields is necessary for forecasting production and projecting the economic impacts of a possible climate change (Lobell *et al.*, 2007). Sun *et al.* (2007) realize a study about the climate influence on crop yields trying to help in the design of policies to reduce the vulnerability related to the climate in Brazil. Some studies analyse the relationship between crop yields and climate change, and Blanc (2012) use regression analysis to quantify the impact of climate on corn in Africa.

Among the studies that have analysed the relationship between climate variables and yields is that of Alexandrov and Hoogenboom (2001), who using the stepwise technique, found that precipitation and maximum temperature anomalies during summer months best explain soya bean yield in Georgia in the United States. Rodríguez Puebla *et al.* (2004) applied a linear regression statistical model

* Correspondence to: M. P. Llano, National Scientific and Technological Research Council (CONICET), Department of Atmospheric and Oceanic Sciences, Faculty of Sciences, University of Buenos Aires, Intendente Güiraldes 2160 (1428) Buenos Aires, Argentina. E-mail: mpllano@at.fcen.uba.ar

to predict wheat yield in Spain using climate variation modes. Sierra and Murphy (1983) explained the variability of maize yield using combinations of various variables in the maize-growing region of Argentina. Penalba *et al.* (2007) analysed the impact of weather on soya bean yield using multivariate analysis. Sarker *et al.* (2003) develop simple models, like stepwise regression, to predict lentil yields as a function of seasonal rainfall and temperature in Near East. Similar study is performed by Bootsma *et al.* (2005) where they determinate relationship between corn and soya beans yields and some agroclimatic variables.

The objective of this study is to summarize the relationships between climatic variables and yields in the three areas where the production of maize and soya bean allows higher export balances. Also, given the wide range of climatic zones where these crops are produced, a division of climatic homogeneous zones is tested, trying to find the most productive areas, they are defined as core areas. The influence of these areas in ranges low, medium and high yield is studied. Finally, this study pretends to characterize climate regimes in selected producing areas and estimate the yield of maize and soya bean as a function of climate variables.

2. Data and methodology

The climate variables used in this study are maximum and minimum monthly mean temperatures and total monthly rainfall. These variables were selected because in terms of general water balance, they provide information regarding incoming and outgoing energy. The data belong to the Climate Prediction Center of the National Center for Environmental Prediction and cover the period from 1979 to 2009. The reference stations are indicated in Table 1, and the geographic location is showed in Figure 1. These reference stations were selected for their quality information and geographic locations, and cover the production areas in Argentina, Brazil and the United States.

Yield is used as a variable to describe crop production for maize and soya bean. It is expressed in units of kilograms per hectare and reflects what happened during each crop cycle for the entire country. That is, it does not address specific locations but is an overall value obtained from the average of all locations where each product is grown. These data are obtained from agricultural statistics of the United Nations' FAO.

The presence of a positive trend is evident in all yield series. Numerous studies have related the existence of this trend with technological advances implemented for crop improvement. Garcia *et al.* (1987) conducted a study linking the levels and variability of maize yields in the United States with technological advances and climate. The same relationship between yields, technology and climate was studied by Tannura *et al.* (2008), in this case for both maize and soya bean in three US states. Hu and Buyanovsky (2003) eliminated the trend present in the yield series because they associate it with advances in technology, while they related interannual variations with climate fluctuations. The trend present in the yield series (5% significance) is removed by fitting a second-degree polynomial (Hazell, 1985). The expressions of these polynomials are:

$$\text{Yield maize Argentina } (t) = 1.7481\, t^2 + 91.418\, t + 2724.4$$
$$\text{Yield maize Brazil } (t) = 1.8338\, t^2 + 22.627\, t + 1635.5$$
$$\text{Yield maize USA } (t) = 2.5077\, t^2 + 41.988\, t + 6322.3$$
$$\text{Yield soya bean Argentina } (t) = 0.1884\, t^2 + 20.062\, t + 1878.7$$
$$\text{Yield soya bean Brazil } (t) = 0.201\, t^2 + 35.299\, t + 1527.6$$
$$\text{Yield soya bean USA } (t) = -0.1928\, t^2 + 36.955\, t + 1865.5 \quad (1)$$

Table 1. Location of the stations used in this study.

Station	Lat.	Lon.	Country	N° map
Resistencia Aero	−27.45	−59.05	Argentina	1
Formosa Aero	−26.2	−58.23	Argentina	2
Villa Maria del Rio Seco	−29.9	−63.68	Argentina	3
Ceres Aero	−29.88	−61.95	Argentina	4
Reconquista Aero	−29.18	−59.67	Argentina	5
Paso de los Libres Aero	−29.68	−57.15	Argentina	6
Córdoba Aero	−31.32	−64.22	Argentina	7
Paraná Aero	−31.78	−60.48	Argentina	8
Concordia Aero	−31.3	−58.02	Argentina	9
Rio Cuarto Aero	−33.08	−64.27	Argentina	10
Marcos Juárez Aero	−32.7	−62.15	Argentina	11
Rosario Aero	−32.92	−60.78	Argentina	12
Junín Aero	−34.55	−60.95	Argentina	13
Santa Rosa Aero	−36.57	−64.27	Argentina	14
Tandil Aero	−37.23	−59.25	Argentina	15
Dolores Aero	−36.35	−57.73	Argentina	16
Tres Arroyos	−38.33	−60.25	Argentina	17
Bahía Blanca Aero	−38.73	−62.18	Argentina	18
Vilhena	−12.73	−60.13	Brazil	1
Bom Jesus da Lapa	−13.27	−43.42	Brazil	2
Vitoria da Conquist	−14.95	−40.88	Brazil	3
Campo Grande	−20.47	−54.67	Brazil	4
Franca	−20.55	−47.43	Brazil	5
Irati	−25.47	−50.63	Brazil	6
Curitiba	−25.52	−49.17	Brazil	7
Paranagua	−25.52	−48.52	Brazil	8
Irai	−27.18	−53.23	Brazil	9
Sao Luiz Gonzaga	−28.4	−55.02	Brazil	10
Santa Maria	−29.7	−53.7	Brazil	11
Birmingham	33.57	−86.75	USA	1
Chattanooga	35.03	−85.2	USA	2
Knoxville	35.82	−83.98	USA	3
Nashville	36.13	−86.68	USA	4
Memphis	35.05	−90	USA	5
Louisville	38.18	−85.73	USA	6
Evansville	38.05	−87.53	USA	7
St. Louis	38.75	−90.37	USA	8
Columbia	38.82	−92.22	USA	9
Peoria	40.67	−89.68	USA	10
South Bend	41.7	−86.32	USA	11
Toledo	41.6	−83.8	USA	12

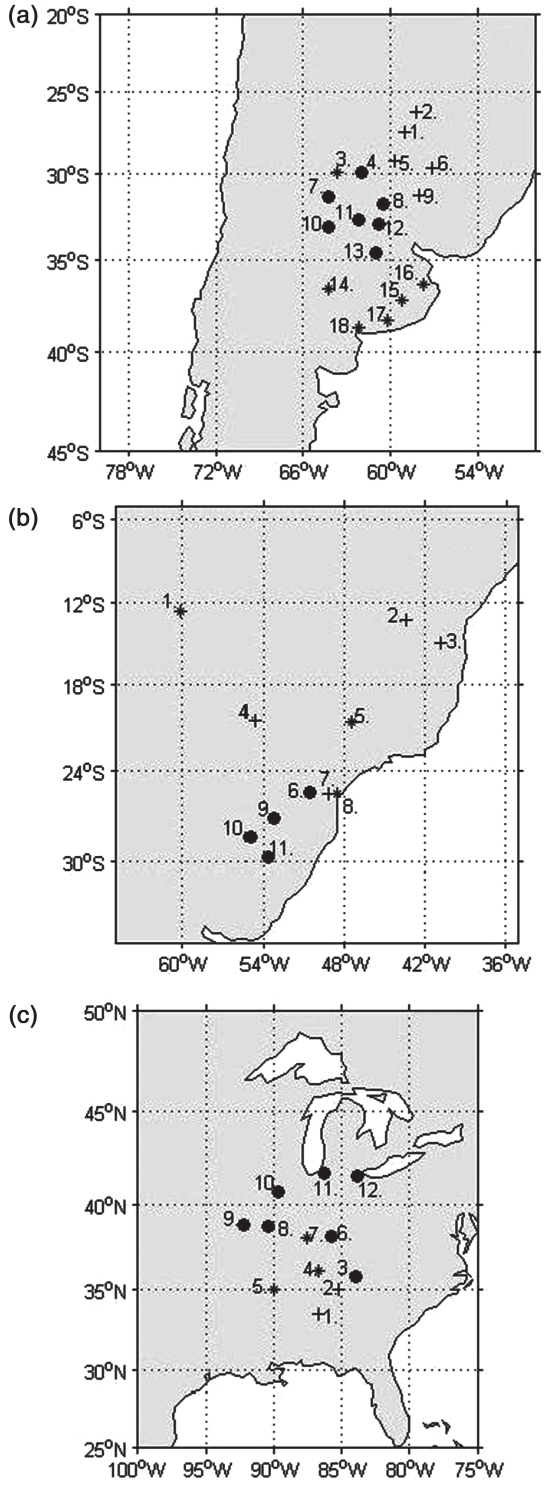

Figure 1. Geographic location of stations used in this study (coded according to Table 1). (a) Argentina, (b) Brazil and (c) United States.

Figure 2 shows the original and detrended maize yield of Argentina. The crop cycle is the time period during which crops evolve through different stages. The main stages are sowing, flowering and harvest. The climate requirements for good development are different in each stage. Soya bean, for example, requires average temperatures above 20 °C during the hottest month of the crop cycle (Pascale and Damario, 2004). In each of the production zones used in this study, the crop cycles have different lengths, and the months they comprise are different (USDA, 1994) (Table 2).

The clustering algorithm K-means [developed by MacQueen (1967)] is widely used in many areas; Zscheischler *et al.* (2012) made a world climate classification, Robertson and Ghil (1998) defined weather regimes of daily precipitation and temperature over the western United States in winter. Naumann and Vargas (2009) used this methodology to make a classification of the thermal structure in different stations of Argentina.

This K-means methodology is used in this study to find reference station groups that have similar characteristics. This method allows for the partitioning of a group of N observations into K groups. These groups are selected such that the sum between the elements of the group and its centroid is minimal. The approach proposed by Hartigan (1975) is used to identify the optimal number of groups. The Hartigan index $H(k)$ relates the squared distance between the elements and the centroids. The optimal number is given for the K that maximize the $H(k)$. The maximum temperature and precipitation during the maize cycle months is used to determine the groups. The maize crop cycle is chosen because, in most areas, the soya bean cycle occurs during the same months. The input matrix for cluster analysis depends on the number stations and campaign months of each country. For example, in Argentina is a matrix with 18 columns (number of stations) × 18 rows (maximum temperature in the ninth first position continuing with the nine values of precipitation – nine months is the duration of maize crop cycle). The number of optimal clusters is three for each country

$$H(k) = (n - k - 1)\,\frac{\mathrm{err}(k) - \mathrm{err}(k+1)}{\mathrm{err}(k)} \tag{2}$$

$$\mathrm{err}(k) = \sum_{i=1}^{k} \sum_{j=1}^{n} d^2\left(X_j, X_{ci}\right) \tag{3}$$

where, d indicates Euclidean distance between the elements X_j and the centroid X_{ci} and n indicates number of meteorological stations.

One of the methodologies used in the study is analysis of variance. The purpose of analysis of variance is to determinate how specific sources of variation contribute to the total variance of a quantity, and to test whether the effect of a particular factor is real or likely to have arisen as a result of random errors (with a significance level of 5%). This method is used to study the variability of rainfall and temperatures when the yield conditions are divided in terciles (low, medium, high) during the different stages.

The stepwise multiple regression method is used to estimate crop yields using climate variables (Draper and Smith, 1981). The method selects the climatic variable most correlated with the yield and finds the first order linear regression equation, and checks if this variable is significant or not (significance level 95%). Then the partial correlation coefficients of all the predictors not in regression are analysed. The highest partial correlation coefficient with the yield is now selected and a second regression

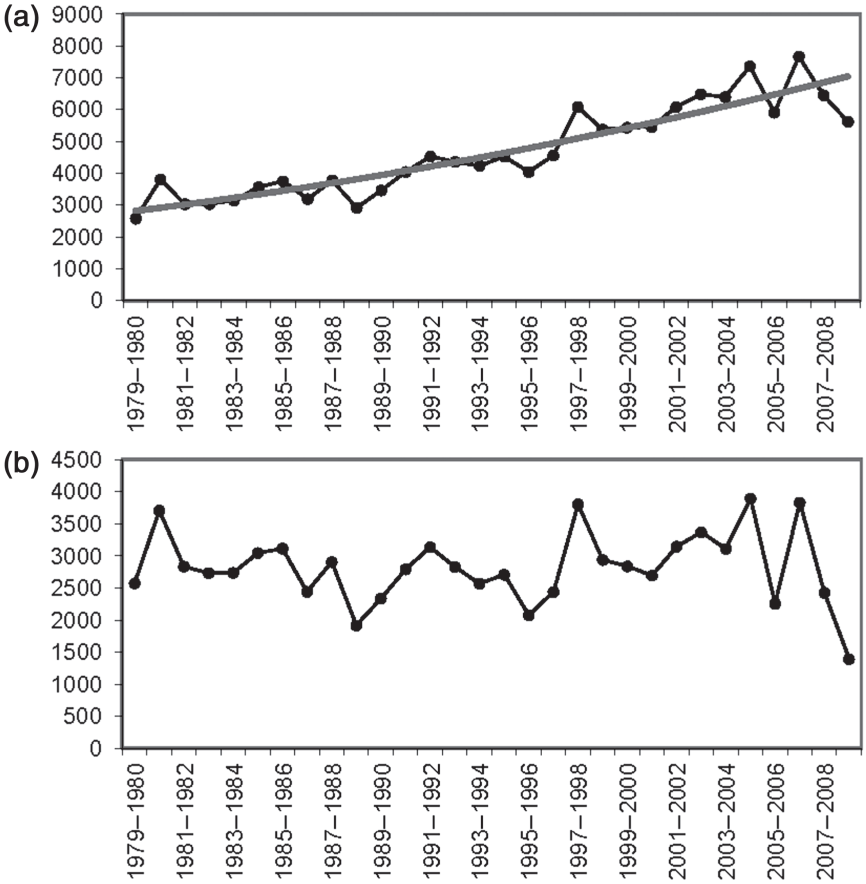

Figure 2. Maize yield in Argentina, (a) original yield with polynomial adjust and (b) detrended yield.

Table 2. Stages of maize and soya bean for different countries.

Maize	January	February	March	April	May	June	July	August	September	October	November	December
Argentina	F	F	H	H	H				S	S	S	F
Brazil	F	F	F	H	H	H					S	S
USA				S	S	F	F	F	H	H	H	
Soya bean	**January**	**February**	**March**	**April**	**May**	**June**	**July**	**August**	**September**	**October**	**November**	**December**
Argentina	F	F	F	H	H						S	S
Brazil	F	F	H	H	H						S	S
USA					S	S	F	F	H	H	H	

F, flowering; H, harvest; S, sowing.

equation with the two climatic variables is fitted. Each predictor variable is retained or rejected according to whether the test is significant or not. This test is carried out at every stage of the stepwise procedure. This methodology is used in works in the climate study area; Rowhani *et al.* (2011) examine the relationship between seasonal climate and maize, sorghum and rice yields in Tanzania. In the same line, Mikova *et al.* (2013) use multiple regression analysis to quantify the maize grain response to climate in Bulgaria.

3. Results

3.1. Climate analysis

The climate variables at each of the reference stations are analysed individually. In Argentina, the maximum temperatures at all stations follow an annual wave. In Figure 3, the Tandil and Formosa stations are shown as examples. In the summer months, the maximum temperature varies between 33 °C at Formosa and 27 °C at Tandil; all other stations have values within this 6 °C range. During winter, the values at all stations vary between 12 and 22 °C. According to the analysis of the extent of the amplitude of maximum temperature, the stations in the northern portion of the study area report a lower amplitude (Formosa barely exceeds 11 °C), while in the southern area the amplitudes increase (Bahía Blanca station registers the highest amplitude, exceeding 17 °C). The rainfall pattern exhibits a dry and a wet season. Stations in the central and northern areas of the region report differences greater than 150 mm between the wettest and driest months. At

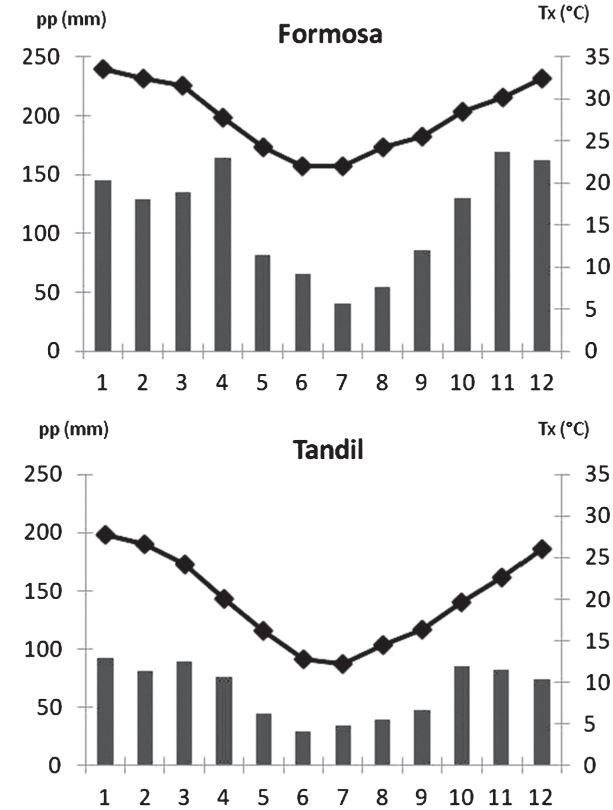

Figure 3. Annual variation of maximum temperature (line with dots) and precipitation (bars) for Formosa station (upper) and Tandil station (lower).

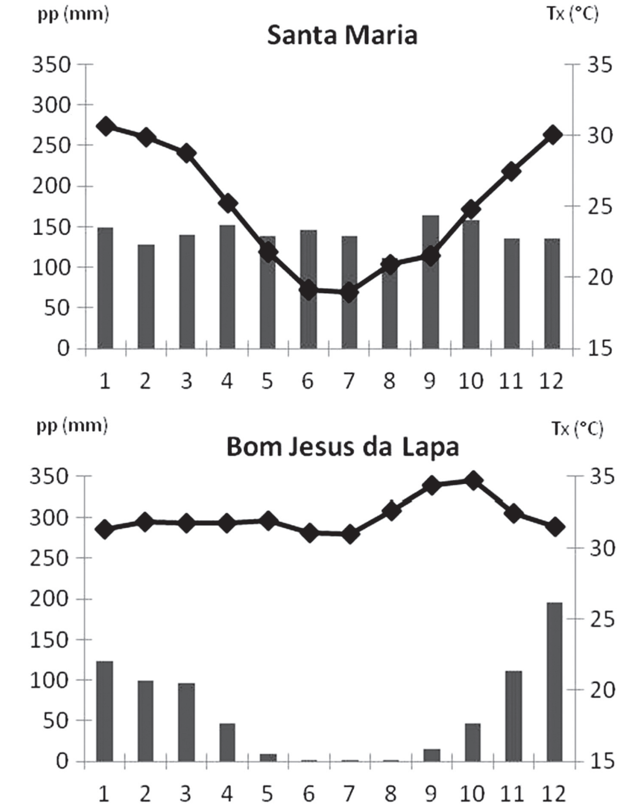

Figure 4. Annual variation of maximum temperature (line with dots) and precipitation (bars) for Santa Maria station (upper) and Bom Jesus da Lapa station (lower).

stations located in the south, the amplitude barely exceeds 50 mm. Northeastern Argentina presents monthly totals of approximately 120–150 mm; these values decrease slowly towards the south, reaching values of 80 mm per month in the south of Córdoba and in Buenos Aires Province.

In Brazil, there are differences in the maximum temperature between stations located in the north of the area of study and those located in the south (Figure 4). Because of their geographical location, the northern stations show the double annual maximum value that is characteristic of inter-tropical zones. The southern stations exhibit the behaviour observed in sub-tropical areas. Bom Jesús da Lapa (in the north) presents the record for the highest temperature, with 34.6 °C in October, and Santa Maria (in the south) holds the record for the minimum, with 19 °C in July. Annual amplitudes of maximum temperature values are between 3 and 12 °C among all stations, north stations show the lowest values. With respect to rainfall, the stations in the north have a clear wet season in December–January–February because of the summer monsoon. In the southern stations, it is possible to distinguish a dry season even though the recorded values are not below 50 mm; in this region, a bimodal regime dominates, with two rainy seasons that occur in March–April and in September–October.

In the United States, the maximum temperatures for all stations show a clear annual wave. In the winter months, there is a 12 °C difference between the northern and southern stations in the study area. In Figure 5, Birmingham station and South Bend station exemplify these differences. During the summer, this gradient is only 3 °C. Total annual rainfall varies between 850 and 1400 mm. These amounts remain uniform throughout the year, and there is no marked dry season, as the absolute minimum is 43 mm and the maximum is 138 mm.

The behaviour of the annual variation of minimum temperatures is not mentioned specifically because of its similarity to maximum temperature. Given these results, it is necessary finding areas of similar behaviour. The joint variability of maximum temperature and precipitation during the maize cycle months is used to determine the groups by K-means methodology.

In Argentina, the patterns of joint variability are similar among the three groups. The main difference observed is in the variation ranges that arise primarily from the geographic location of stations and their climate regime. Figure 6(a) shows the first group (called 'north') which comprises the northern stations (Resistencia, Formosa, Reconquista, Paso de los Libres and Concordia, –- represented by '+' in Figure 1(a)). The sowing months show a cumulative increase in precipitation (with a doubling in its value), and the same occurs with the maximum temperature. The three flowering months show similar rainfall amounts (120 mm), and during this period, the maximum temperatures reach their highest recorded levels (33 °C), as expected for this time of year. During harvest,

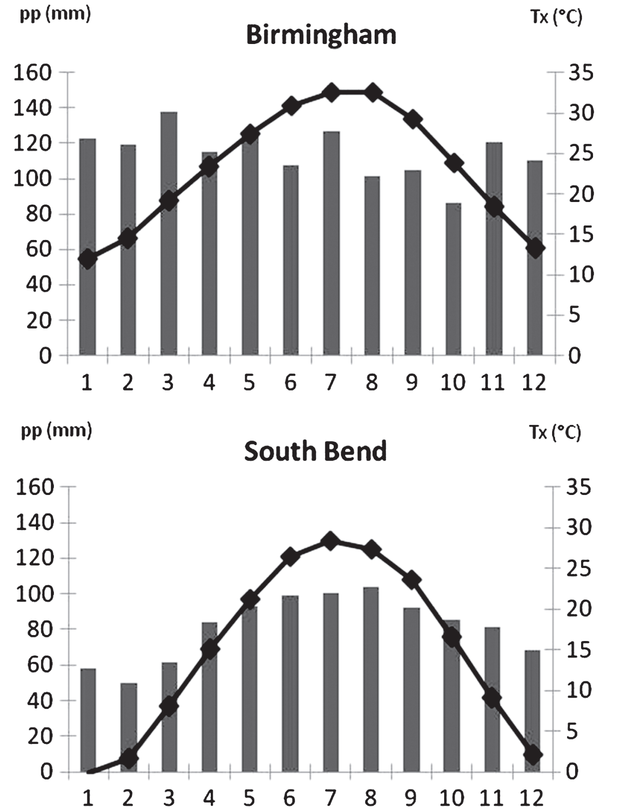

Figure 5. Annual variation of maximum temperature (line with dots) and precipitation (bars) for Birmingham station (upper) and South Bend station (lower).

temperatures decrease with an average of 3 or 4 °C per month. Rainfall during the first 2 months exhibits levels above those observed during the flowering stage, but at the end of the crop cycle, rainfall decreases to the levels observed at the beginning of the flowering stage.

The second group (called 'centre', Figure 6(b)) comprises the stations located in the so-called Humid Pampas (Ceres, Córdoba, Paraná, Río Cuarto, Marcos Juárez, Rosario and Junín – 'o' symbol in Figure 1(a)). The precipitation increases in each month during the sowing. In the first month of the flowering stage, total rainfall peaks (140 mm). In the remaining 2 months of this stage, precipitation decreases and reaches the second relative maximum (130 mm) in the first month of harvest, thus generating a particular shape for the curve. In the remaining months of harvest, accumulated rainfall decreases, again reaching a value of approximately 40 mm.

The third group (called 'south', Figure 6(c)) comprises the stations located in the south of the study area (Santa Rosa, Tandil, Dolores, Tres Arroyos, Bahía Blanca), to which Villa María is added ('*' symbol in Figure 1(a)). During the sowing months, accumulated precipitation increases; throughout the flowering stage, the rainfall records are stable (70 mm); and during harvest, decreases between 30 and 40 mm are observed each month. With respect to temperature, variations during the crop cycle show no striking features, as their behaviour is consistent with this time of year.

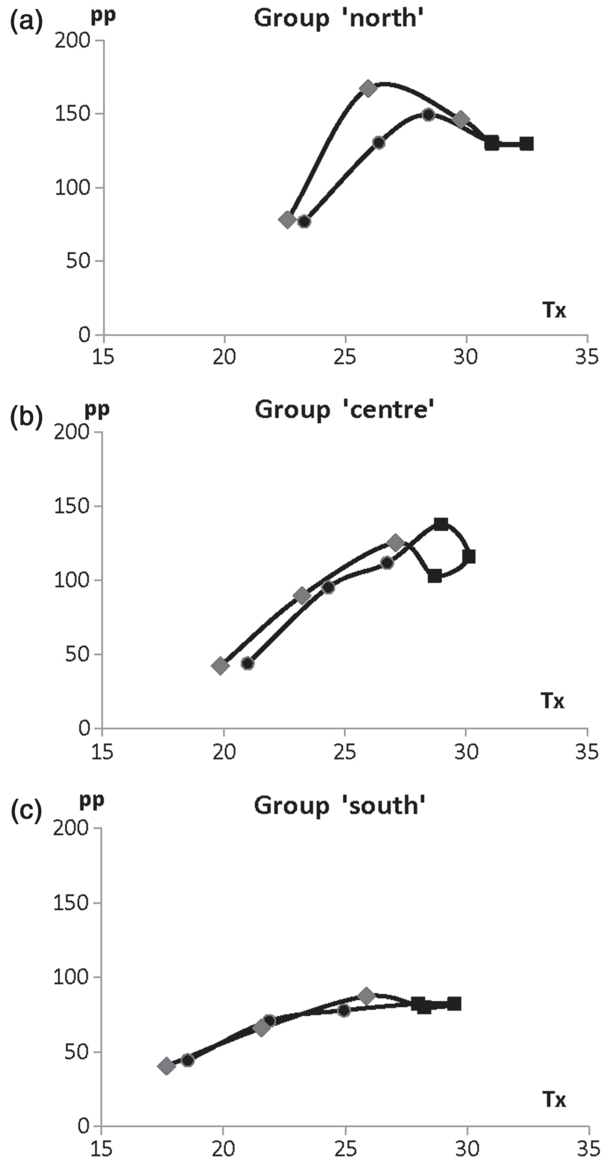

Figure 6. Joint evolution of maximum temperature (°C) and precipitation (mm) during maize cycle, for the three centroid groups of Argentina. Sowing stage is represented by black circles, flowering stage by black quadrate and harvest stage by grey diamond.

For Brazil, the K-means method reveals three different groups, which are exhibited by three patterns. As shown in Figure 7(a), the first group (called 'south') presents similar behaviour during the sowing months, with a maximum temperature of 28 °C and 140 mm of rainfall. During the flowering stage, there is very little difference in the total rainfall, with a maximum of 160 mm, and temperatures are near 29 °C. In the three harvest months, rainfall does not vary, while maximum temperatures decrease. The stations associated with this pattern are Iratí, Irai, Sao Luiz Gonzaga and Santa María, which are located in the main producing zone ('o' symbol in Figure 1(b)).

The stations associated with the pattern for the second group (called 'centre', Figure 7(b)) are Vilhena, Franca and Paranagua ('*' symbol in Figure 1(b)). This pattern is characteristic, given the low variability in temperatures. During

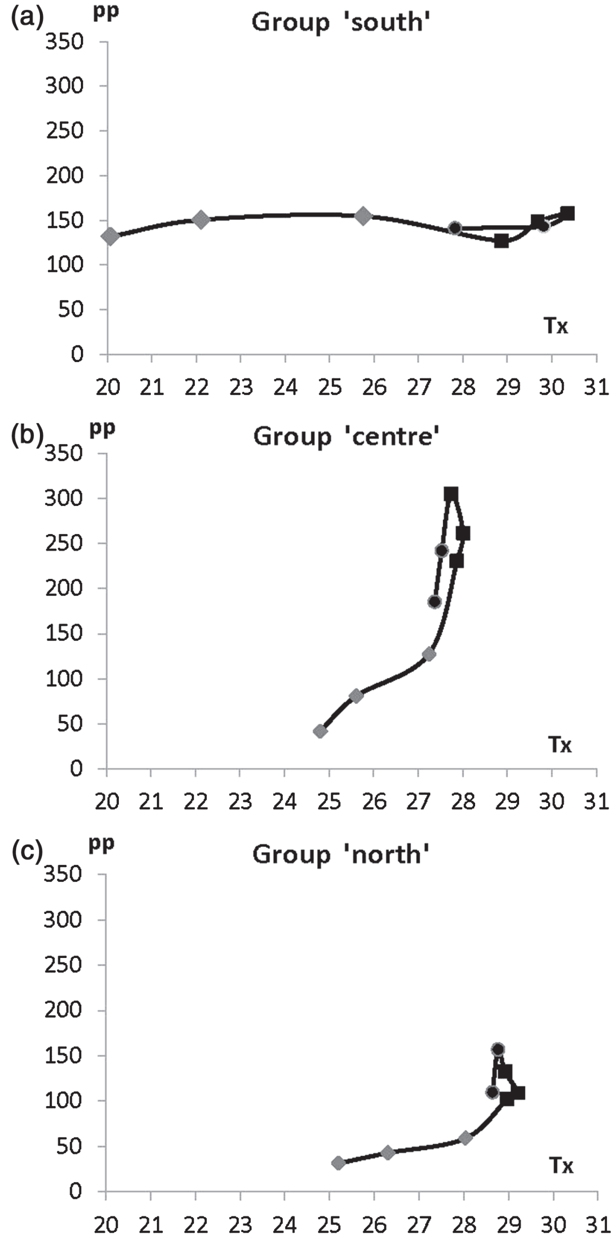

Figure 7. Joint evolution of maximum temperature (°C) and precipitation (mm) during maize cycle, for the three centroid groups of Brazil. Sowing stage is represented by black circles, flowering stage by black quadrate and harvest stage by grey diamond.

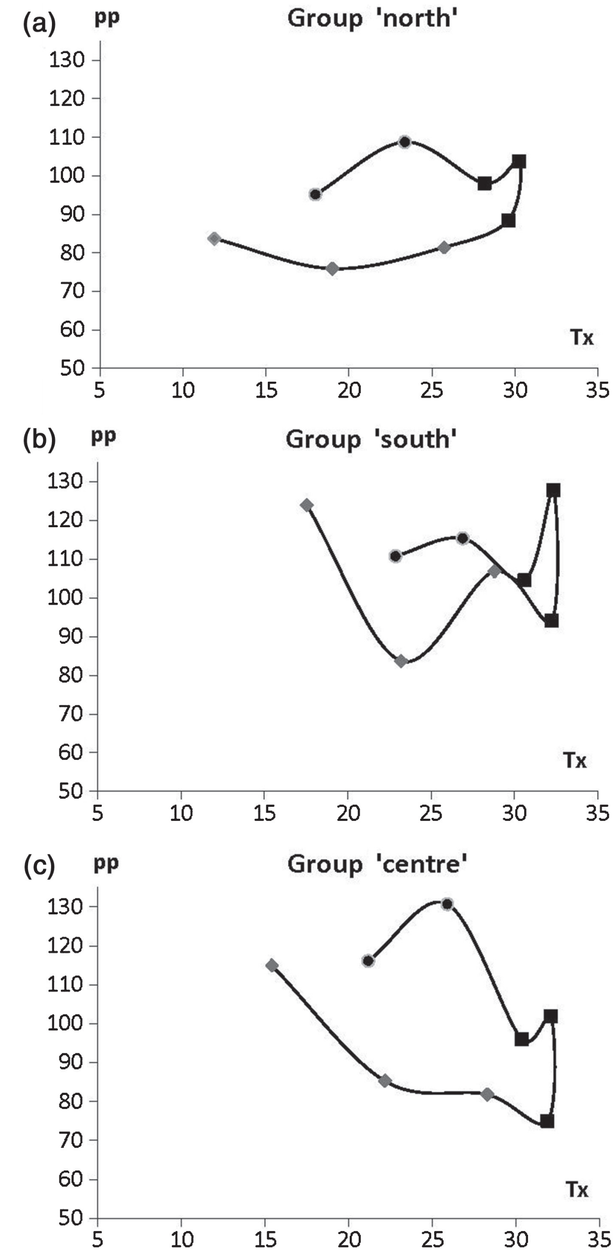

Figure 8. Joint evolution of maximum temperature (°C) and precipitation (mm) during maize cycle, for the three centroid groups of United States. Sowing stage is represented by black circles, flowering stage by black quadrate and harvest stage by grey diamond.

the sowing months, precipitation increases by 60 mm, and temperatures remain near 27 °C. The first month of the flowering stage shows the highest cumulative rainfall, 300 mm, which then decreases; the maximum temperature, approximately 28 °C, does not vary over all 3 months. At harvest, there is a significant decrease in accumulated precipitation. Maximum temperatures also decrease, reaching 24 °C in the last month of the crop cycle.

Finally, the third group (called 'north', Figure 7(c)) includes the Bom Jesús da Lapa, Vitoria da Conquist, Campo Grande and Curitiba stations ('+' symbol in Figure 1(b)). During the sowing and flowering stages, the behaviour of both variables is very similar during all months. High temperatures are near 28 °C, and rainfall ranges between 100 and 160 mm. In the three harvest months, temperatures decrease to 25 °C. Rainfall exhibits a similar pattern, as recorded values decrease gradually, reaching 30 mm in the last month of the crop cycle.

In the United States, optimal division also results in three groups. As shown in Figure 8(a), the first group (called 'north') includes the largest number of stations: Knoxville, Louisville, Saint Louis, Columbia, Peoria, South Bend and Toledo (in Figure 1(c) represented by 'o' symbol). Joint variability increases for both variables during the sowing stage. During the flowering stage, the maximum temperature is approximately 29 °C, and rainfall ranges between 90 and 105 mm. During harvest,

Table 3. Results of variance analysis for each K-means grouping in each country.

	Precipitation						Maximum temperature						Minimum temperature					
	North		Centre		South		North		Centre		South		North		Centre		South	
Argentina	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S
Sowing	R	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR
Flowering	nR	nR	R	nR	nR	nR	nR	nR	R	R	nR	nR	nR	nR	nR	nR	nR	nR
Harvest	R	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR
	Precipitation						Maximum temperature						Minimum temperature					
	South		Centre		North		South		Centre		North		South		Centre		North	
Brazil	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S
Sowing	R	nR	nR	nR	R	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR
Flowering	R	R	nR	nR	nR	nR	nR	R	nR	nR	nR	nR	nR	nR	R	nR	nR	nR
Harvest	nR	nR	nR	nR	nR	nR	nR	R	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR
	Precipitation						Maximum temperature						Minimum temperature					
	North		South		Centre		North		South		Centre		North		South		Centre	
USA	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S
Sowing	nR	nR	nR	nR	nR	nR	nR	nR	nR	R	nR	nR	nR	R	nR	nR	nR	R
Flowering	nR	nR	R	nR	R	nR	R	R	R	R	R	nR	R	nR	nR	nR	R	nR
Harvest	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR	nR

The letter R (nR) indicates the refusing (not refusing) of null hypothesis (not difference in climatic variables depending yield categories). In the different stages of maize (M) and soya bean (S) crop cycle.

maximum temperatures reach 15 °C, and precipitation reaches values close to 80 mm (with a small increase in the last month of the crop cycle).

The second group (called 'south', Figure 8(b)) shows month-to-month fluctuations in precipitation. During the sowing months, there is an increase in temperature and accumulated precipitation. The precipitation presents a minimum during the first and last month of the flowering stage, and the maximum for the crop cycle (130 mm) occurs in the second month of this stage. There is not much variability in temperature, they remain at approximately 32 °C. Finally, during harvest, the maximum temperature decreases from 28 to 17 °C. Rainfall decreases between the first and second month (105–80 mm); the last month contains the second relative maximum of the crop cycle, 125 mm. The stations in this group are Birmingham and Chattanooga ('+' symbol in Figure 1(c)).

In the last group (called 'centre'), the temperature behaviour is very similar to that observed in the other two groups; for this last group, values range between 15 and 32 °C (Figure 8(c)). Rainfall shows an absolute maximum in the second month of the sowing stage (130 mm); during the flowering stage, accumulated rainfall decreases to between 75 and 100 mm each month. Finally, during harvest, rainfall gradually increases, reaching a value of 115 mm (the third relative maximum). The stations associated with this group are Nashville, Memphis and Evansville ('*' symbol in Figure 1(c)).

3.2. Regional variability components

The analysis of variance is used to estimate the parts played by yield categories (high, medium and low) in the total variance of the series of climatic variables. The null hypothesis is that there is no difference in climatic variables depending on the yield categories. A regional average for climatic variables is realized in each zone (defined by K-means). The results for each crop stage are displayed in Table 3.

For Argentina, when rainfall is the variable analysed, the null hypothesis is not rejected in most cases, indicating that accumulated precipitation has no influence on yield. However, there are particular cases in which the hypothesis is rejected which are of interest. The group 'centre' during flowering, for example, indicates that at this stage, accumulated precipitation affects the yield. For temperature, only the maximum temperatures during flowering have some relationship with the categories of both crop yields.

For Brazil, in most cases, the hypothesis that there is no difference in total rainfall between yield categories is not rejected. However, this assumption is rejected for the 'south' group, for both the maize sowing and flowering stages and for the soya bean flowering stage, showing the important role of precipitation in the final yield. When the variable analysed is extreme temperature, maximum temperatures show a relationship with yield categories during the soya bean flowering and harvest stages for group 'south', while minimum temperatures only have a relationship with maize yields during the flowering stage in group 'centre'.

For the United States, in general, the hypothesis that precipitation is similar in all categories of yield is not rejected, except in special cases in which there is a category of yields during the flowering stage that shows values of accumulated precipitation that are different from the rest. This result could indicate that the precipitation

Table 4. Coefficients of correlation between maximum temperatures and precipitation.

	Maize			Soya bean		
R	Sowing	Flowering	Harvest	Sowing	Flowering	Harvest
Argentina	**−0.38**	**−0.37**	**−0.38**	**−0.56**	−0.33	−0.25
Brazil	**−0.55**	**−0.42**	−0.36	**−0.55**	−0.27	**−0.46**
USA	**−0.40**	**−0.50**	−0.21	**−0.44**	**−0.44**	−0.21

The significant coefficients (95%) in bold letter.

produced at this stage affects the yield. Maximum temperatures during the flowering stage of both crops and for all groups show a relationship with the categories of yield, as the null hypothesis is rejected. Minimum temperatures show fewer such cases, as they occur mainly during the maize flowering stage and the soya bean sowing stage.

Specific areas (hereinafter 'core areas') are selected in each country in accordance with major crop areas of USDA Book (1994) and based on the results of the climatology analysis and variance analysis. In Argentina and Brazil, the group with more null hypothesis rejected is selected. For the United States, the results of variance analysis does not show a clear principal zone, for this reason the group with the largest number of stations is selected, which includes the stations of the main producer crop area. For Argentina, the core area is formed by the stations in group 'centre'; for Brazil, the core area is formed by stations in group 'south'; and the group 'north' is the core area for the United States. These areas present the major production; this has special interest in this study.

3.3. Estimation of maize and soya bean yields

As another point in the analysis of the climate-crop relationship, the stepwise regression technique is used to analyse the weight of the climate variables measured in the stages on the final yield of crops in each core areas. The parameters used in this methodology represent regional average climatic variables (precipitation and maximum temperatures averaged for each stage).

The availability of this yield information is in a country scale. This information is underestimating the yields in core areas. Total country value is an average between core area and other regions with lower yields (according to FAO statistics). Monthly climatic variables in the core area are assumed as representatives of the thermal condition and precipitation throughout all the production area in each country.

The existing linear relationship between precipitation and monthly maximum temperature is studied for each of the different stages of development of the crop. In the cases where the relationship between temperature and precipitation is significant, the effect is filtered out and the analysis is performed with the part of the maximum temperatures that is not related to precipitation. Table 4 shows the correlation coefficients between the regional average climatic variables.

The theoretical models for both crops in the three regions are as follows:

$$\begin{gathered}
\mathrm{Y_M}\ \text{Argentina} = 1590.9 + 2.59\ pp_\mathrm{F} - 271.54\ T'x_\mathrm{F} \\
\mathrm{Y_M}\ \text{Brazil} = 1379.74 + 0.59\ pp_\mathrm{F} + 0.79\ pp_\mathrm{S} \\
-\ 0.45\ pp_\mathrm{H} \\
\mathrm{Y_M}\ \text{USA} = 2215.91 - 462.99\ T'x_\mathrm{F} + 146.42\ T'x_\mathrm{S} \\
+165.16\ Tx_\mathrm{H} \\
\mathrm{Y_S}\ \text{Argentina} = 4209.88 + 1.51\ pp_\mathrm{F} - 89.22\ pp_\mathrm{S} \\
-106.71\ T'x_\mathrm{F} \\
\mathrm{Y_S}\ \text{Brazil} = 1092.8 + 1.44\ pp_\mathrm{F} - 96.51\ T'x_\mathrm{S} \\
+\ 82.12\ T'x_\mathrm{H} \\
\mathrm{Y_S}\ \text{USA} = 1685.46 - 82.07\ T'x_\mathrm{F} \qquad (4)
\end{gathered}$$

where Y represents the estimated yield, subscript M indicates the maize crop, and subscript S the soya bean crop. The variables involved are precipitation (*pp*), maximum temperatures (*Tx*) and maximum temperatures not explained by precipitation (*T'x*) for the three developmental stages: sowing (S), flowering (F) and harvest (H). For example, ppS represents the time series of regional average of precipitation (previously averaged among the months) when sowing is produced. Figures 7 and 8 show the anomalies of original yield (OY) and estimated yield (EY) and the standard deviation (SD) of OY.

For Argentina (Figure 9(a)), the theoretical model for maize yield is significant, explaining 49% of the variance of the original yield. This is also directly related to precipitation during the flowering stage and inversely related to the maximum temperature that is not related to precipitation during the flowering stage. This model indicates the importance of flowering because the values of the climate variables during this stage define the level of yield at the end of the crop cycle. During the study period, the EY is near to the original, but not in the last years where the values are separated to the mean value.

For Brazil (Figure 9(b)), the theoretical model of maize yield involves more climate variables, which account for 44% of the variance of the OY. Rainfall during the sowing and flowering stage (directly) and at harvest (indirectly) is significantly related to yield. The OY anomalies are between the standard values at the beginning of the period. The EY accompanies these values. During the last years, a major variability in the original yield produces that the theoretical model does not represents the values.

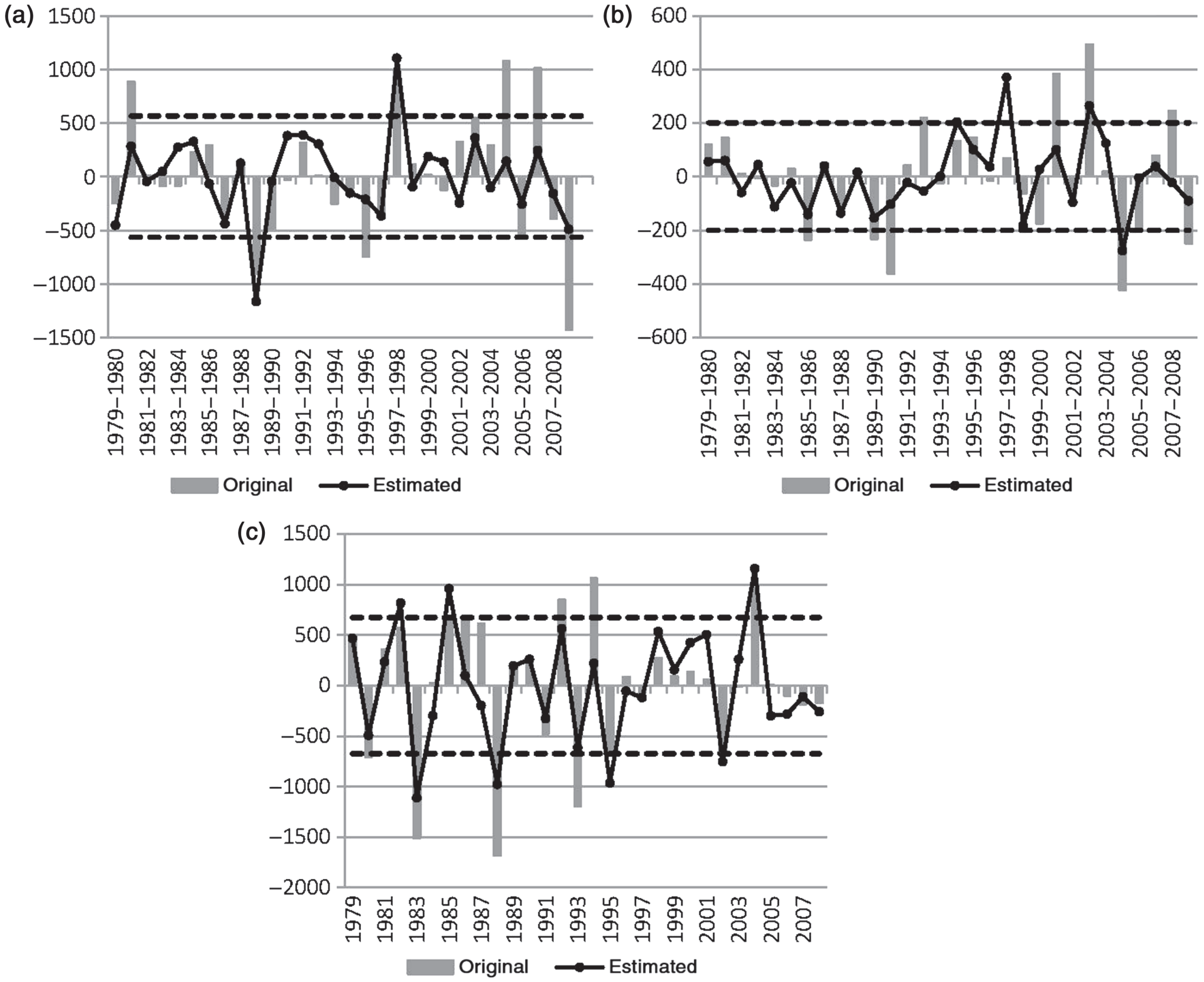

Figure 9. Anomalies of detrended original yield of maize in three producing regions (grey bars), anomalies of estimated yield (black line) and one standard deviation of original yield (dotted black line). (a) Argentina, (b) Brazil and (c) United States.

For the United States (Figure 9(c)), the model explains 70% of the variance of the OY. The climate variables that show a significant relationship (5%) are maximum temperatures not explained by precipitation during the sowing (direct) and flowering (indirect) stages. The model is completed with the harvest maximum temperature. During the complete period, the two yields are similar including the extreme events.

For the cultivation of soya bean in Argentina (Figure 10(a)), the obtained theoretical model explains 66% of the variance of the OY. The variables in this model that show a significant relationship are precipitation during the flowering stage in a direct form, and maximum temperature at flowering and precipitation during the sowing, with an indirect relationship. The original yield variability is represented by the estimated yield along the complete period.

The theoretical model for Brazil explains 51% of the variance of the OY (Figure 10(b)). The model is composed three significant variables. Rainfall during the flowering stage shows a positive relationship. The part of maximum temperatures that is not explained by precipitation during the sowing stage exhibits an inverse relationship, while during harvest it shows a direct relationship. The variability around the mean value is bigger in the second part of the period. This can be one of the reasons for the differences between estimated and original yields.

The model for soya bean for the United States explains 37% of the variance of the OY (Figure 10(c)). The model contains only the part of maximum temperatures that is not related to precipitation during the flowering stage. The original yield presents variability that is not represented by the estimated yield, mainly in the extreme values.

The six models show a closer relationship between the yield and the climatic variables during the flowering. As shown in Table 5, in all theoretical models, the adjustment is statistically significant at the 5% level. The root mean square error is less than the SD of the original yield, indicating the usefulness of representing annual yields through climate variables.

3.4. Climate characteristics in campaign with extreme yields

The climatic study of temperatures and precipitation and the variance analysis allow the grouping of stations and the selection of core areas in each country. The climatic

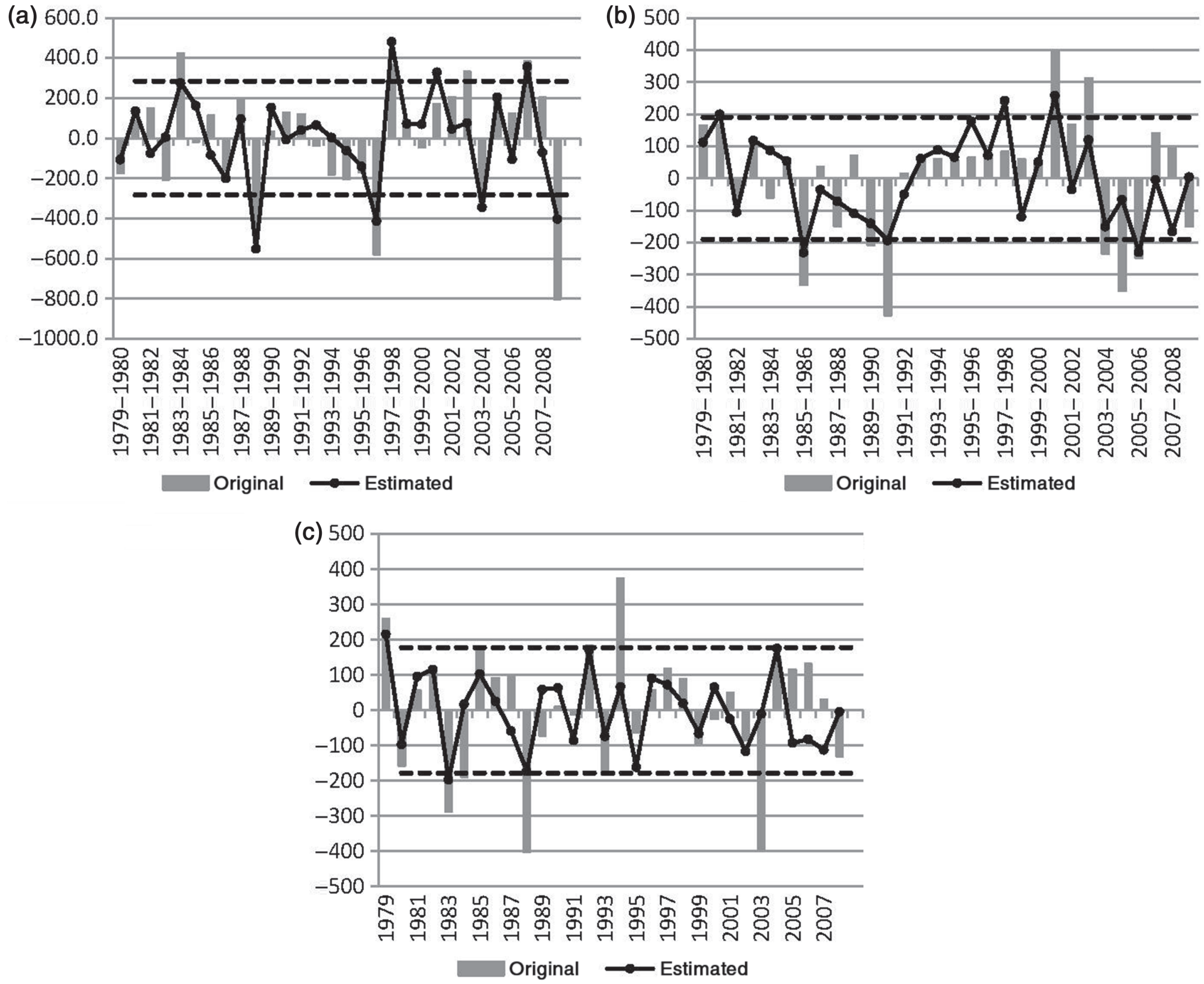

Figure 10. Anomalies of detrended original yield of soya bean in three producing regions (grey bars), anomalies of estimated yield (black line) and one standard deviation of original yield (dotted black line). (a) Argentina, (b) Brazil and (c) United States.

Table 5. Statistics of the regression model.

Maize	R	R^2	df	RMSE	SD	Mean value
Argentina	0.70	0.49	2	397.0	564.9	2817.6
Brazil	0.66	0.44	3	147.6	200.4	1660.0
USA	0.84	0.70	3	358.9	673.8	6374.1
Soya bean	R	R^2	df	RMSE	SD	Mean value
Argentina	0.81	0.66	3	161.5	283.2	1898.9
Brazil	0.72	0.51	3	130.5	190.1	1563.0
USA	0.60	0.37	1	139.0	177.9	1903.2

R, coefficient of correlation; R^2, coefficient of determination; df, degrees of freedom; RMSE, root mean square error (in kg ha^{-1}); SD, standard deviation (in kg ha^{-1}) and mean value of yields (in kg ha^{-1}).

variables in these areas show a good relationship with the crop yields. The stepwise linear regression allows to determine the relationship between yield and climatic variables during the period 1979–2009. The estimated yields and the results of variance analysis represent mean situations, but there are special campaigns where the results are not adjusted to typical events, it may be because of the extreme behaviour of climatic variables.

In Figures 9 and 10, particular years where the theoretical model overestimates or underestimates the original yield are detected where a significant departure from the mean value is observed. For example, maize in Argentina at the end of the period presents three special campaigns. The 2008–2009 campaign shows a yield lower than the mean value (2 SD – SD-). This value can be attributed to precipitation during flowering and harvest, which are in the first and second QUINTILE of the distribution, respectively. The harvest temperature has a special performance (3 SD more than its mean value). Campaigns 2004–2005 and 2006–2007 with yields bigger than the average (2 SD) are characterized by more precipitation during the flowering (fourth and fifth QUINTILE, respectively). The value of the theoretical model is greater than the average but does not represent the extreme conditions. The behaviour of these campaigns can be compared with 1997–1998 campaign, where the original yield is 2 SD above the average, and it can be estimated by the theoretical model. In these years, the precipitation is in the fifth QUINTILE.

For soya bean in Argentina, the theoretical model overestimates the original yield in the last year in 3 SD. During this campaign, temperatures values are higher than normal in all stages while precipitation during sowing is in the first QUINTILE and normal during flowering. Campaigns where yield is above its average (1 SD) show normal or slightly lower temperatures; the precipitations during sowing and flowering are in the fifth QUINTILE.

In Brazil, the 1997–1998 maize campaign has a normal yield which is overestimated by the theoretical model. In this year, precipitation during flowering is in the fifth QUINTILE, 880 mm (doubling the mean value). Campaigns 2000–2001 and 2002–2003 with a high yield (2 SD more than the mean value) are underestimated by the theoretical model mainly on account of the precipitation variability in the different stages. The 1993–1994 campaign shows normal values in the original and estimated yields, temperatures are lower than the average and precipitation is high during sowing (fourth QUINTILE), but normal in flowering and harvest stages.

The soya bean in Brazil exhibits some particular campaigns. For example, 1990–1991 shows a low yield (2 SD) with normal temperatures but few precipitations during flowering (in the first QUINTILE). An opposite situation is presented by campaigns 2000–2001 and 2002–2003, where for high yields (2 SD more than the average), the flowering precipitations are in the fifth QUINTILE and in the other stages the values are normal.

Maize in the United States also presents same special campaigns; e.g. in 1988, the OY is 2 SD lower than the average. In this year, the temperatures are normal but the sowing and flowering precipitations are deficient (in the first QUINTILE of the distribution). In 1994 and 2004, the OY are bigger than the average, in the first case the EY overestimates the original yield while in 2004 the OY is underestimated. The principal differences are the temperature values during the stages.

Soya bean in the United States presents special campaigns like year 1988, where the OY is 2 SD below the normal value, mainly because of the low precipitation during sowing (which is not included in the theoretical model) and the high temperature in the same stage. Another situation is the 1994 campaign, where OY is 2 SD greater than the average, with precipitation in these stages in the first or second QUINTILE and the temperature is above the normal value. The 2003 campaign presents an excess in the precipitation and cold temperatures, and this situation can be the responsible of the EY value.

In search of coherence between the three countries, it is possible to notice that in Argentina and Brazil precipitation is the principal discriminator of yield, whereas in the United States temperature takes that role. The reason of that could be the differences in the precipitation regimes between the United States and Argentina and Brazil.

4. Conclusions

One motivation of this study is the necessity to understand the relationship between the yields of maize and soya bean and different climatic variables, mainly because Argentina, Brazil and the United States are the largest producers as well as three of the leading exporting countries. This means that the crises in the availability to export in these countries affect other usual importing countries. These crises are in general associated with extreme climatic variability which may be identified by extreme temperatures and precipitations. However, this is an approximation because yields are also affected by other factors, like economic variables or ground water storage.

One of the difficulties representing yields in the regions is given by the wide range of variability in the climatic variables and the fact that crops can adapt with different yield grades. It is for these reasons that considering the joint variability between temperature and precipitation three zones are discriminated in each country. With these divisions and using the analysis of variance, core areas are selected where the climate–yield relationship is the closest, consistent with the highest production areas documented by FAO.

In Argentina, the core area yield is proportional to the country scale yield, which value is lower than the core area yield (Argentine Ministry of Agriculture, Livestock and Fisheries – http://www.siia.gov.ar/). The yield in the core area can be considered approximately equal to a constant by the yield throughout the production area. This can cause that the lineal regression determines a relationship with a quantity proportional to the yield in the core area.

The yield fitted model shows that the climatic variables measured during flowering stage have the greatest influence on the final crop yield. The estimated relationships are linear, involving different variables in each country. Maize yield in Argentina is related to precipitation and maximum temperature, while the fitted soya bean yield is best represented by the sowing and flowering precipitation and maximum temperature, this is in concordance with the study performed by Penalba *et al.* (2007) that found a relationship within soyabean yield and climatic variables in central-eastern Argentina. In Brazil, for maize, the precipitation in the three stages is related to the yield. Franchito *et al.* (2010) have found that more precipitation during rainy station (flowering months) is related with more corn yield in San Pablo state. The maximum temperature in sowing, and harvest and flowering precipitation are related to the soya bean yield. In the United States, maximum temperatures in the three stages are related to the maize yield. While for soya bean only the flowering maximum temperature presents a relationship. This result is consistent with Alexandrov and Hoogenboom (2001) who found a good relationship between the maximum temperature, precipitation and yield of soya bean in the state of Georgia (USA). The usefullness of these models is to get a good synthesis of the yields in function of climatic variables. The constants of the models could be used to typify temporal and spatial climate variability and the yields variability in maize and soya bean production areas.

This study pretends to analyse the different parameters needed to describe the yield. However, it is necessary to add daily variables in the regression model because the use temperature and precipitation stage averages in statistical approaches cannot show other aspects of sub-seasonal variations. These variations such as dry spells or heat waves can be critical to the crop growth (Porter and Semenov, 2005). The mentioned three countries are analysed given the interest in knowing the different characteristics of the relationship between crops and climate that are involved in foreign trade.

Acknowledgements

We thank anonymous referees for their valuable comments and critical reading of the article. This research was sponsored by projects UBA 2002130300014BA and FONCyT PICT 2012-1507.

References

Alexandrov V, Hoogenboom G. 2001. Climate variation and crop production in Georgia, USA during the twentieth century. *Clim. Res.* **17**(1): 33–43.

Blanc E. 2012. The impact of climate change on crop yields in Sub-Saharan Africa. *Am. J. Clim. Change* **1**: 1–13.

Bootsma A, Gameda S, McKenney D. 2005. Potential impacts of climate change on corn, soybeans and barley yields in Atlantic Canada. *Can. J. Soil Sci.* **85**(2): 345–357.

Draper N, Smith H. 1981. *Applied Regression Analysis*, 2nd edn. John Wiley and Sons: New York, NY.

Franchito S, Brahmamanda Rao V, Gan M, Santo C. 2010. Onset and end of the rainy season and corn yields in Sao Paulo, Brazil. *Geofís. Int.* **49**: 69–76.

Garcia P, Offutt SE, Pinar M, Changnon SA. 1987. Corn yield behavior: effects of technological advance and weather conditions. *J. Clim. Appl. Meteorol.* **26**: 1092–1102.

Hartigan J. 1975. *Clustering Algorithms*. John Wiley and Sons: New York, NY.

Hazell PBR. 1985. Sources of increased variability in world cereal production since the 1960s. *J. Agric. Econ* **36**: 145–159.

Hu Q, Buyanovsky G. 2003. Climate effects on corn yield in Missouri. *J. Appl. Meteorol.* **42**: 1626–1635.

Kucharik C, Serbin S. 2008. Impacts of past and future changes in climate and atmospheric CO_2 on Wisconsin agriculture. Executive Summary October 2008. Environmental and Economic Research and Development Program, 6 pp, Wisconsin.

Lobell DB, Cahill KN, Field CB. 2007. Historical effects of temperature and precipitation on California crop yields. *Clim. Change* **81**: 187–203.

MacQueen J. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Volume I: Statistics*, Le Cam LM, Neyman J (eds). University of California Press: Berkeley and Los Angeles, CA, 281–297.

Mikova A, Alexandrova P, Dimitrov I. 2013. Maize grain yield response to n fertilization, climate and hybrids. *Bulg. J. Agric. Sci.* **19**: 454–460.

Naumann G, Vargas W. 2009. Changes in the predictability of daily thermal structure in southern South America using information theory. *Geophys. Res. Lett.* **36**: L09704, doi: 10.1029/2009GL037576.

Pascale A, Damario E. 2004. *Bioclimatología Agrícola y Agroclimatología*, 1st edn. Facultad de Agronomía, Universidad de Buenos Aires, Ciudad de Buenos Aires, Argentina.

Penalba O, Bettolli M, Vargas W. 2007. The impact of climate variability on soybean yields in Argentina. Multivariate regression. *Meteorol. Appl.* **14**: 3–14.

Porter JR, Semenov MA. 2005. Crop responses to climatic variation. *Philos. Trans. R. Soc. B: Biol. Sci.* **360**: 2021–2035.

Robertson A, Ghil M. 1998. Large-scale weather regimes and local climate over the Western United States. *J. Clim.* **12**: 1796–1813.

Rodríguez Puebla C, Encinas A, Frías M. 2004. Relaciones entre la variabilidad climática y la productividad de trigo en España. *Física de la Tierra* **16**: 1–13.

Rowhani P, Lobell D, Linderman M, Ramankutty N. 2011. Climate variability and crop production in Tanzania. *Agric. For. Meteorol.* **151**: 449–460.

Sarker A, Erskine W, Singh M. 2003. Regression models for lentil seed and straw yields in Near East. *Agric. For. Meteorol.* **116**: 61–72.

Sierra E, Murphy G. 1983. Variabilidad del rendimiento del maíz en la región maicera argentina. *Meteorológica* **14**: 73–85.

Sonka ST, Changnon SA, Hofing S. 1998. Assessing climate information. Use in agribusiness. Part II: decision experiments to estimate economic value. *J Clim* **1**: 766–774.

Sun L, Li H, Ward M, Moncunill D. 2007. Climate variability and corn yields in semiarid Ceará, Brazil. *J. Appl. Meteorol. Climatol.* **46**: 226–240.

Tannura MA, Irwin SH, Good DL. 2008. Weather, technology, and corn and soybean yields in the U.S. corn belt. Marketing and Outlook Research. Report 2008-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Illinois.

Thompson LM. 1969. Weather and technology in the production of corn in the US corn belt. *Agron. J.* **61**: 453–456.

USDA, United States Department of Agriculture. 1994. Major world crop areas and climatic profiles. In *Agricultural HandBook*, Washington, D.C.

Zscheischler J, Mahecha MD, Harmeling S. 2012. Climate classifications: the value of unsupervised clustering. *Procedia Comput. Sci.* **9**: 897–906.

