METEOROLOGICAL APPLICATIONS
Meteorol. Appl. **14**: 3–14 (2007)

(www.interscience.wiley.com) DOI: 10.1002/met.1

The impact of climate variability on soybean yields in Argentina. Multivariate regression

Olga C. Penalba,[1]* M. Laura Bettolli[1,2] and Walter M. Vargas[1,2]
[1] *Departamento de Ciencias de la Atmósfera y los Océanos, FCEN, Universidad de Buenos Aires, Argentina*
[2] *Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina*

ABSTRACT: Climate variability is examined and discussed in this work, emphasizing its influence over the fluctuation of soybean yield in the Pampas (central-eastern Argentina). Monthly data of rainfall, maximum and minimum temperatures, thermal range and seasonal rainfall were analysed jointly with the soybean yield in the period 1973-2000. Low-frequency variability was significant only in the minimum temperature during November in almost all the stations. This situation is favourable to the crop since during this month, seed germination, a growth stage sensitive to low temperatures, takes place. In the crop's core production region, 72% of the series of soybean yield presented a positive trend. Except in years with extreme rainfall situations, interannual variability of the soybean yield is in phase with the seasonal rainfall interannual variability. During these years, losses in the soybean crop occurred, with yield negative anomalies greater than one standard deviation. Soybean yield showed spatial coherence at the local scale, except in the crop's core zone. The association between each climate variable and yield did not show a defined regional pattern. Summer high temperature and rainfall excesses during the period of maturity and harvest have the greatest negative impact on the crop, whilst higher minimum temperatures during the growing season favour high yields. The joint effect of climate variables over yield was studied with multivariate statistical models, assuming that the effect of other factors (such as soil, technology, pests) is contained in the residuals. The regression models represent the estimates of the yield satisfactorily (high percentage of explained variance) and can be used to assess expected anomalies of mean soybean yield for a particular year. However, the predictor variables of the yield depend on the region.

KEY WORDS climate variability; soybean yield; statistical models; Argentina

Received 17 January 2006; Revised 28 November 2006; Accepted 28 November 2006

1. Introduction

Cultivation of grain crops constitutes the most important agricultural activity in Argentina. They have been, and continue to be, subject to a wide range of study in the areas of economics, agronomics and meteorology. The agricultural activity involves a broad set of decision-making in which a variety of factors have significant influence. In this context, climate is a source of variability and risk, causing, in some situations, a negative impact on agricultural activities. Therefore, it is necessary to assess the level of influence of temporal and spatial climate variability on crop yields. This evaluation is hindered on account of the complex quantification of the technological component present intrinsically in this activity.

Cultivation of the soybean crop in Argentina began in the early 1960s with 10 000 sown hectares (Bolsa de Cereales, 1979). Today, over 12 000 000 hectares are cultivated, constituting one of the most important crops in Argentinian agriculture. The influence of climatic variables on soybean crop yield has been studied generally in controlled experiments at reduced space-time scales. In Argentina, few studies have examined yield behaviour over extended periods, probably because such data series are scarce. Calviño and Sadras (1999) studied the response of the soybean yield in terms of the interaction of precipitation, soil depth and handling practices in two grower-managed fields located in Buenos Aires province. They found a significant positive correlation between yield and soil water availability in deep soil during January and February. Ravelo *et al.* (1983) developed a biometeorological model based on the effect of the daily maximum and minimum temperatures, together with the length of day in the development of soybean crop at one station in Buenos Aires. This model allowed forecasting dates of occurrence of soybean phenological phases.

Undoubtedly, these studies in controlled experiments contribute significantly to the knowledge of the interaction between soybean and the environment. However, at the level of larger time and space scales, where the climate factors cannot be controlled, the analysis of the climate-agriculture relationship would allow better planning of agricultural activities, with the purpose of preventing or mitigating the negative impacts and taking

---

* Correspondence to: Olga C. Penalba, Departamento de Ciencias de la Atmósfera y los Océanos, Universidad de Buenos Aires, Ciudad Universitaria – Pab II, 1428-Buenos Aires, Argentina.
E-mail: penalba@at1.fcen.uba.ar

advantage of the positive ones. Among others, the influence of climate upon the different phenological stages of the crop was studied by Pascale and Escales (1971, 1975); Pascale and Murphy (1975). They studied the bioclimatic requirements of the soybean crop in Argentina by analysing different sub-periods of the vegetative cycle (sowing/flowering and flowering/maturity) from 14 years of experimental data. They found that within the sowing/flowering stage, the most influential variables are the air temperature and the length of day, while in the flowering/maturity stage they are the air temperature and the edaphic and atmospheric humidity. The areas most suitable for soybean cultivation turned out to be the eastern Argentine region (Figure 1): the south of Santa Fe Province, north of Buenos Aires Province, the centre-southeast of Córdoba Province and the west of Entre Ríos Province (Pascale and Rodríguez, 1977). More recently, Hurtado *et al.* (2001) studied the variations in soybean yields in relation to soil water storage in the Pampas region. They found that low values of yields were related to situations of scarce water storage in the soil during November, December and January, while high values correspond to a wider range of hydric situations. Boullón (2002) showed that in the Argentine Pampas region, water stress, hail and frosts are the weather phenomena that can produce a reduction from severe to extreme in the crop yield, in different growing stages.

Research studies where the relationship between climatic variables and soybean yield is quantified through statistical models at a regional scale are scarce. In Argentina, Minetti and Lamelas (1995) studied the response of soybean yield to climatic variability by means of a method of multiple regression in San Miguel de Tucumán (north-west of Argentina). They found that December rainfall and February mean thermal range are the variables that most relate to the crop yield. They also observed that during the summer months (December, January and February) the need for water is greater, whereas in the four-month term of January, February, March and April, soybean needs more air moisture. In Georgia, United States, Alexandrov and Hoogenboom (2001) found that the variables that best explain soybean yield are the July and August precipitation anomalies and the September maximum temperature anomalies using a stepwise regression model. Meanwhile, in a region of the northeast United States the climatic variables associated with this crop are monthly rainfall and mean temperature during July and August (Huff and Neill, 1982).

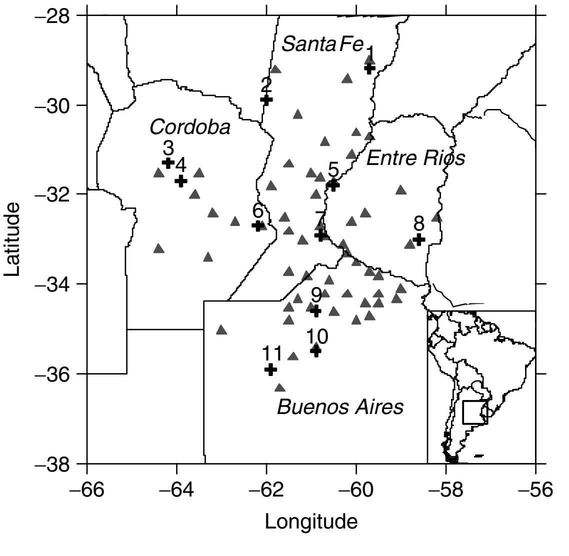

Figure 1. Location of the meteorological stations (crosses) and provincial departments (triangles–the agricultural data coordinate corresponds to the geographical centre of each department).

Climate variability and change caused by natural processes, as well as anthropogenic factors, are major and important environmental issues that will affect the world during the twenty-first century. World agriculture, whether in developing or developed countries, remains very dependent on climate resources. Therefore, the impact of climate variability on agricultural production is important at local, regional, national, and global scales. The main goal of this work is to assess the impacts of climate variability on agriculture in the agricultural region par excellence of Argentina, the Pampas region. Specific objectives were to: (1) analyse temporal and spatial variability of the (monthly and seasonal) climatic variables and soybean yield; (2) analyse the relationship between seasonal rainfall and yield; (3) objectively diagnose the relationship between soybean yield and climatic variables; and (4) analyse the regional stability or variability of this relation.

The following section describes the data-set used, the methodology, the climatic regime and soybean yield behaviour in the study region. Results in Section 3 report the temporal variability of the climatic variables and the soybean yield in the core production region of Argentina; the relationship between climatic variables and soybean yield; and the possible climate variables diagnostic capability per yield amount. The conclusions are summarised in Section 4.

2. Data and methodology

2.1. Climatological data

Climatological data came from 53 rain gauges that provided daily rainfall data and from 64 meteorological stations, which supplied daily data on rainfall and maximum and minimum temperatures, operated by Servicio Meteorológico Nacional, Argentina. The region and the period of study were governed by the availability of agricultural data and therefore, restricted to the north of parallel 40 °S.

After evaluating daily weather data for erroneous and missing values, only the stations with 10% or less missing data were further considered in this study. A quality control was performed on monthly rainfall and maximum and minimum temperatures. Statistical analyses were performed so that mean values, standard deviations and distributions were compared for every station and month using the Student-*t*, Chi-Square and Fisher tests ($a = 5\%$, Höel, 1964; Panofsky and Brier, 1965). This procedure was applied according to the guidelines set in the control guide of quality surface climatic data published by the World Meteorological Organization (WMO; Abbott, 1986) within the World Climate Data program. Figure 1 and Table I show the location of the 11 meteorological stations that passed the quality control. Data used in this paper are monthly rainfall, monthly mean maximum and minimum temperatures and monthly mean thermal range, as a measure of air humidity, in the period (1973–2000). The months analysed were those considered within the mean growing cycle (November–May of the next year) including the pre-sowing month of October (next section).

Because rainfall is the most significant climatic variable in the determination of crop yield (Giorda and Baigorri, 1997) and soybean production in Argentina takes place mostly without irrigation, the rainfall totals series (hereafter referred to as 'seasonal rainfall') in the mean growing cycle were also analysed.

2.2. Soybean data

This work focuses on yield (estimated as the ratio of total production to area harvested) as an indicator of a crop's vulnerability to climate variability. On the basis of the analyses of Technical Reports of the sown and harvested zones available at the Secretaría de Agricultura, Ganadería, Pesca y Alimentación de la Nación, (SAGPyA, 2000) the soybean mean growing season of the analysed zone was defined from November (sowing) to May (harvest). These were the months taken into consideration given that both the sown and harvested zones surpass 70% during this period. October is considered the pre-sowing period. In this study, a growing cycle is noted by the year in which a crop was sown, even though harvest takes place in the following calendar year (e.g. the 1982–1983 cropping season is noted as cropping season 1982). SAGPyA supplied the yield series of 301 provincial departments. Those series that showed inconsistency in the soybean information (no communication between the sown and harvested areas, yield and crop production) were eliminated. The length of the records of these series is variable. In order to obtain series with the largest possible register length (to provide statistically stable results) and a proper spatial covering, the 1973 growing season turned out to be the optimum starting season. Finally 58 departments were used in this study in the period 1973–1999 (Figure 1, triangles).

Table I. Number, name, latitude and longitude of the meteorological stations used in the study.

No.	Meteorological stations	Latitude (*S*)	Longitude (*W*)
1	Reconquista	29°11′	59°42′
2	Ceres Aero	29°53′	61°57′
3	Córdoba Aero	31°19′	64°13′
4	Pilar	31°40′	63°53′
5	Paraná	31°47′	60°29′
6	M Juárez Aero	32°42′	62°09′
7	Rosario	32°55′	60°47′
8	Gualeguaychú Aero	33°00′	58°37′
9	Junín Aero	34°33′	60°55′
10	9 de Julio	35°27′	60°53′
11	Pehuajó	35°52′	61°55′

2.3. Methodology

Low-frequency variability of temporal series was studied based on the analysis of the linear trend. The significance of the linear slope was considered at 5% (Höel, 1964). The Pearson's first moment correlation was applied to study: (1) the association between the yield and the climatic variables, and (2) the spatial coherence of variable yield (Wilks, 1995).

Seasonal rainfall and soybean yield were standardized by subtracting their long-term mean and by dividing it by their standard deviation. This procedure enables comparisons among stations with different mean values and standard deviations and facilitates the interpretation of spatial patterns (Brooks and Carruthers, 1953).

In order to estimate the joint effect of the climatic variables in the final yield of the crop, a stepwise multilinear regression model was used (Draper and Smith, 1981). This statistical model selects the predictor variables according to their levels of importance and only if they produce a significant contribution to the variance accounted for by the regression. Each predictor variable is evaluated for its individual significance level before being included in the equation and, with each addition, each variable within the equation is then evaluated for its significance as part of the model. A variable is included in the equation if it is significant at the 95% level and is retained if it is significant at the 99% confidence level. The degrees of freedom and the significance of the regression depend on the number of the predictor climate variables and the length of the series under study. Taking into account this statistical limitation and in order to analyse the responses and the stability of the models, several adjustments with different predictor variables were applied.

3. Results, further analysis techniques and discussion

3.1. Climatic aspects and soybean yield behaviour

The region under study is located in the humid Argentine Pampas. Distribution of extreme monthly mean (minimum and maximum) temperatures shows a predominantly south-north gradient, affected only by the latitude factor. The minimum temperature presents spatial gradients slightly more intense than those of the maximum temperature, with variations ranging from 13 to 16 °C in November, from 16 to 20 °C in January and from 11 to 15 °C in April (months considered representative of the start of the sowing, flowering and harvesting times respectively, for the soybean crop). Maximum temperature variation ranges are within the order of: 26–28 °C in November, 30–32 °C in January and 23 to 25 °C in April. Mean thermal amplitude follows the behaviour of the fields of extreme temperatures analysed previously, with minimum values to the north–northeast. Monthly mean rainfall fields are the most heterogeneous variable presenting different gradient directions according to the time of year. Mean rainfall values vary from 110–130 mm in November, to 70–140 mm in April. Spatial distribution of seasonal rainfall mean values shows a gradient with a southwest to northeast predominant direction coinciding with the annual rainfall pattern (Hoffmann, 1975; Schwerdtfeger, 1976).

In relation to the soybean yield, the mean pattern shows a core higher than 2100 kg ha^{-1} at the north of Buenos Aires, south of Santa Fe and southeast of Cordoba (Figure 2(a)). The largest standard deviations of the yield series are observed to the central-northwest of the region, decreasing towards the east (Figure 2(b)).

3.2. Variability analysis

3.2.1. *Seasonal rainfall*

In the initial analysis, interannual variations and low-frequency trends of seasonal rainfall were investigated for the 11 meteorological stations. Figure 3(a) shows the interannual variations of the standardized seasonal rainfall and the three-year running averages for 4 stations whose results are representative of the regional behaviour. Rosario (Station 7) and Marcos Juárez (Station 6) are located in the core soybean zone presenting the highest mean yields and Paraná (Station 5) and Junín (Station 9) are located to the north and south of the core area respectively. The variation in seasonal rainfall did not show a statistically significant linear trend in the analysed period for any of the 11 stations despite the annual rainfall increments observed during the last decades in central–eastern Argentina (Hoffmann *et al.*, 1987; Castañeda and Barros, 1994).

Seasonal rainfall in the region varied considerably from year to year during the study period. If the attention is focused on the periods when the average seasonal precipitation (the solid curve in Figure 3) has positive values to its long-term mean, two such periods are presented: 1975 to 1981 and 1989 to 1994, in Marcos Juárez (Station 6) and Rosario (Station 7). Between these wet periods, there is one with negative precipitation values (1982–1988). These interannual variabilities appeared in the other stations but with different periodicities. The extreme case is observed at Paraná station (Station 5), which presents the longest dry period (1982–1996) in agreement with Penalba and Vargas (2004) in their research on the interannual variability of the annual rainfall. The occurrence of negative and positive extreme standardized rainfall (values lower than -1 and higher than 1 respectively) showed up in isolated years. The highest probability of occurrence was shown in Junín (Station 9) with 6 years out of 27 showing negative extreme standardized rainfall. Spatial coherence in the occurrence of negative extreme values (years 1973, 1984 and 1996) is higher than in the case of positive extreme values. However, the intensity of these positive extreme values is higher (standardized values close to 2) and in these situations, their spatial coherence is also higher. For example, 1980 shows positive extreme values affecting practically the whole region, with the exception of the stations located in the south of the studied zone (Junín in Figure 3(a)). The same results were observed by Penalba and Vargas (2001).

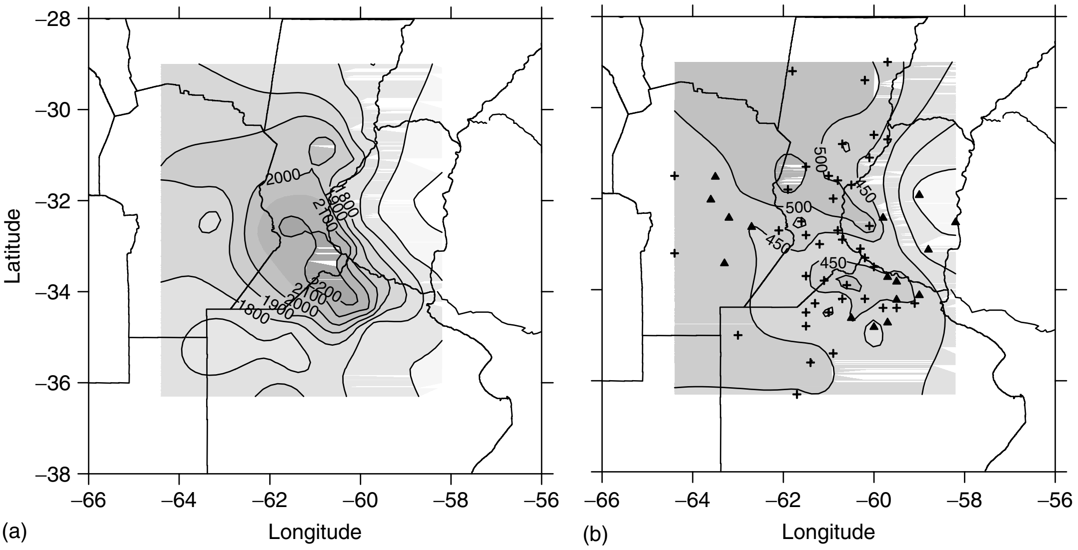

Figure 2. Soybean yield (a) mean field (kg ha^{-1}) and (b) standard deviation (kg ha^{-1}) Crosses: yield series with significant linear trend at the 95% confidence level. Triangles: yield series with no linear trend.

DOI: 10.1002/met

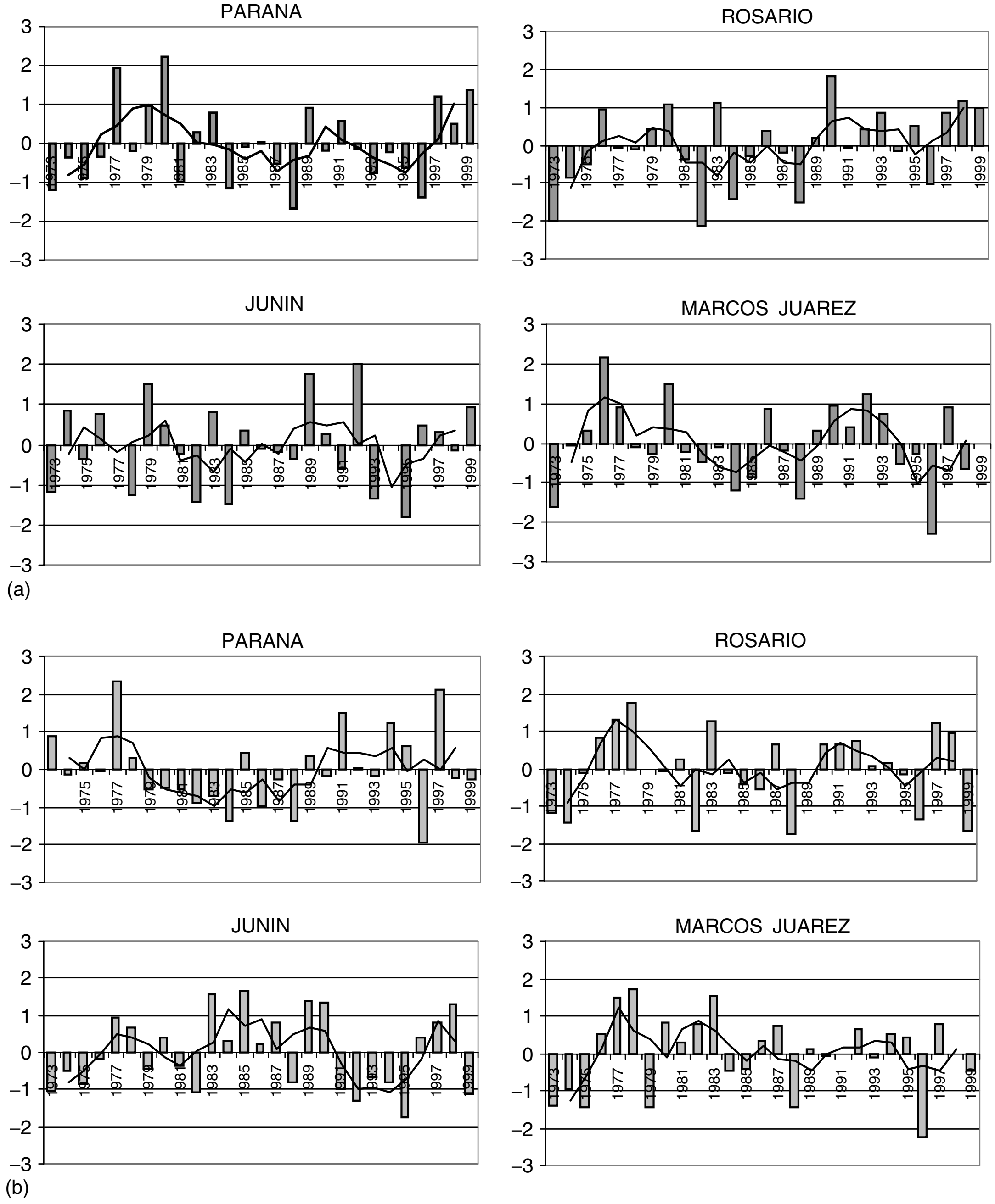

Figure 3. (a) Standardized anomalies of seasonal rainfall (bars) and three-year running averages (solid curve) for Paraná, Rosario, Junín and Marcos Juárez, (b) Idem Figure 3a for soybean yield.

3.2.2. Monthly climatic information

As a part of the climate variability study, monthly rainfall data, monthly mean maximum and minimum temperatures and monthly mean thermal range values were also analysed. For each variable and month (October to May) the linear trend was analysed (Table II) (95% significance level, Wilks, 1995). In general, positive significant trends are observed in the monthly mean minimum temperature series, indicating a regional increase in this temperature (Table II). This result is coincident with those of Barrucand and Rusticucci (2001) and Easterling *et al.* (1997) who analysed longer periods than the one used in this work. November was the month that presented a positive significant linear trend in most of the stations (Pehuajó (Station 11), 9 de Julio (Station 10), Ceres (Station 2) and Córdoba (Station 3) being the exception), a favourable situation for the soybean (high soil temperature during germination) (da Mota, 1978; Yao, 1981) (Table II). The progressive increase or decrease of the monthly mean maximum temperature, has only been shown in isolated months and stations, although this variable presents significant regional trends so far in the twentieth century (Easterling *et al.*, 1997; Barrucand and Rusticucci, 2001). Monthly rainfall has only presented a significant trend in three stations (Junín (Station 9), Rosario (Station 7) and Gualeguaychú (Station 8)), without coincidence in

DOI: 10.1002/met

Table II. Month and indication of significant linear trend, per meteorological station and variable (Min T: minimum temperature, Max T: maximum temperature, ΔT: termal range).

Station	Min T	Max T	Rainfall	ΔT
Pehuajó		(−)Apr		
Junín	(+)Nov		(−)Jan	
9 de Julio	(−)Feb	(−)Apr		
Ceres				
Reconquista	(+)Nov (+)Dec			
Rosario	(+)Oct (+)Nov (+)Dec		(+)Apr	(−)Apr
Paraná	(+)Oct (+)Nov (+)Dec (+)Apr (+)Jun			(−)Oct (−)Nov (−)Dec (−)Apr (−)May (−)Jun
Gualeguaychú	(+)Nov	(+)Nov	(+)Dec	(−)Apr
Córdoba	(−)Jan (−)Feb	(−)Oct (−)Jan (−)Apr (−)May		(−)Apr
Pilar	(+)Nov			
Marcos Juárez	(+)Nov			

the months of occurrence. This could be because of the length of the series since Penalba and Vargas (1996) and Rusticucci and Penalba (2000) found positive significant trends in the monthly rainfall, mainly from December to March (Table II). Monthly mean thermal amplitude showed significant negative linear trends, mainly during April, in only a few stations (4 out of 11).

Both the medium- and short-term variabilities are of interest in this work, and so the linear trend was filtered. From here onwards the analysis is carried out with the anomalies of the climatic variables (regarding the linear trend or the mean, according to the case).

3.2.3. *Agricultural information*

The temporal variations of soybean yield present a large regional variability. Out of the 58 yield series, 42 presented a significant positive linear trend (confidence level 95%) in the analysed period (Figure 2(b)). These series are located in the same areas where the yield has the highest standard deviations (Santa Fe and north of Buenos Aires province). This maximum variability (high standard deviation values) might be due to to the effect of the linear trend. On filtering the trend, the standard deviation field of the series does not show such high variabilities (not shown). Agricultural yield data typically have an upward low-frequency trend because of technological improvements in crop genetics and management techniques (Hall *et al*., 1992). However, the increased variability of yields could be either a function of the heightened sensitivity of technology to weather, or to temporal increases in weather variability (Garcia *et al*., 1987). Many authors in the development of their studies have removed the trend of the yield of several crops, amongst them the soybean, attributing it mainly to technological development (Huff and Neill, 1982; Magrin *et al*., 1998; Podestá *et al*., 1999; Alexandrov and Hoogenboom, 2001; Hurtado *et al*., 2001). Conversely, Minetti and Lamelas (1995) and Krepper *et al.* (1998) associated the trend of the soybean yields with the low-frequency variability in the climatic variables, mainly the summer rainfall (October to April).

However, given that both the medium- and short-term variabilities are of interest in this work, the analysis was performed with the anomalies of the yield (regarding the linear trend or the mean, according to the case).

The interannual variations of the standardized soybean yield (anomalies divided by the standard deviation) and the 3-year running averages for four departments, which are the nearest to the meteorological stations shown in Figure 3(a), are presented in Figure 3(b). In general, it is observed that periods with positive yield and negative anomalies show a sectorized spatial coherence, a result that will be confirmed in the next section. At the same time, interannual variability of soybean yields is generally accompanied by interannual variability in seasonal precipitation. However, exceptions are observed, such as the case of Junín (Station 9), wherein periods with positive anomalies in yield (1983–1987) present negative precipitation anomalies. Furthermore, it is interesting to observe that in the years with negative rainfall extreme anomalies, losses in soybean crops were severe. This direct association does not happen in the analysis of the positive rainfall values (years 1980 and 1997 in Figures 3(a) and (b)). Statistically, these results confirm that excess rainfall during the growing cycle can have different impacts on the yield according to the stages in which they happen, causing disease, floods and failed harvest (Podestá *et al*., 1999; Hurtado *et al*., 2001).

3.2.4. *Spatial coherence in soybean yield*

The spatial structure of the soybean yield field was analysed by calculating the simple correlation among the 58 series of yield anomalies. Figure 4 shows the correlation fields of five departments (located at the north, south, east, west and centre of the study region). The highest spatial coherence corresponds to the department located at the central zone. This significant correlation field extends across almost the whole considered region. Correlation patterns centred in the western and eastern zones of this region are more localized and restricted to each respective zone. The lowest spatial coherence is the one presented by the correlation fields located at the north and south border zones. These results indicate the low spatial coherence of the yield, except for the pockets located in the soybean core region. Therefore, the variable yield is representative in sub-regions. This

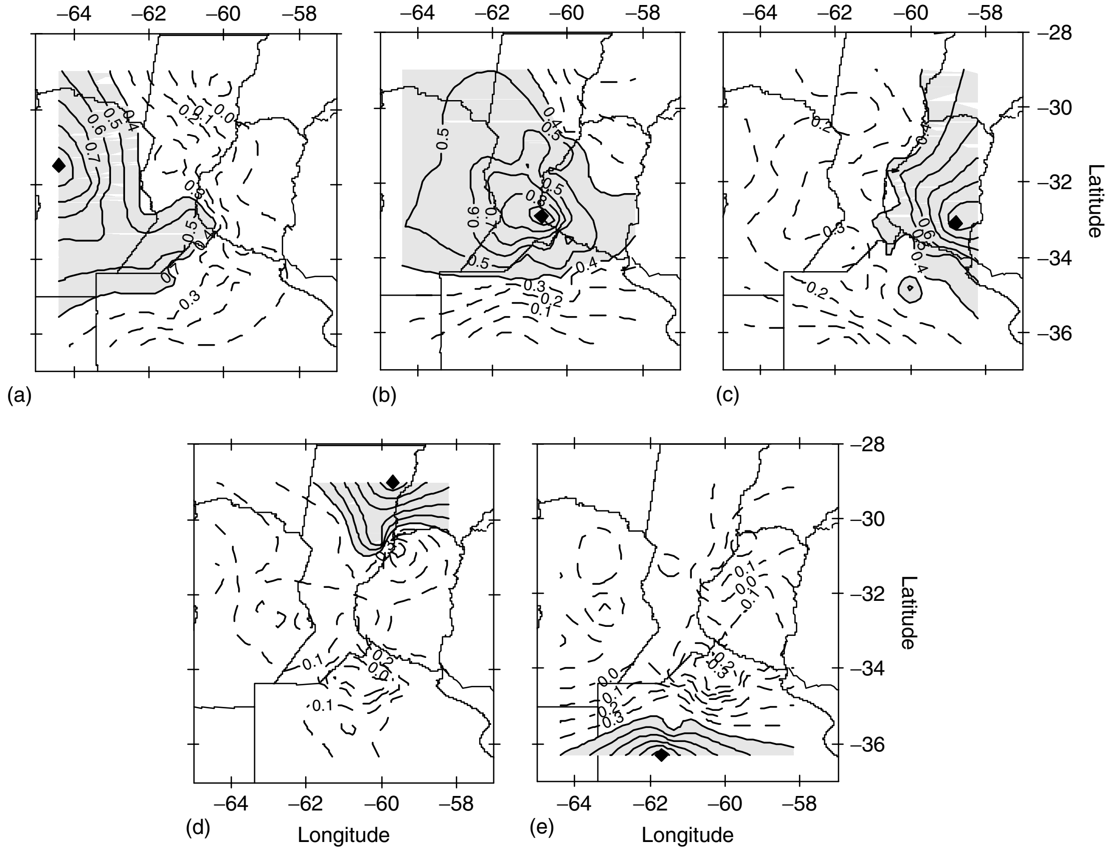

Figure 4. Correlation fields of soybean yield centred on points west (a), centre (b), east (c), north (d) and south (e) (diamonds). Shaded area: correlation coefficients significant at the 95% confidence level.

suggests that regional estimations of soybean yield have to be analysed carefully on account of the low spatial coherence (for example, a real average of the yield would not be representative for the whole region).

3.3. Climatic variables and soybean yield association

Soil and climatic conditions in the study area are suitable for cultivation of soybean. However, sometimes rainfall and temperature extremes occur during the critical period of soybean development. Initially, in order to evaluate the degree of association between climate variability and soybean yield, the simple correlation between each one of the monthly climatic anomalies and yield anomalies were calculated when there was a meteorological station in the province department (Table III).

In general, a marked regional behaviour of the relationship between the variables is not observed in spite of a defined pattern in the indicators of the correlation, whether or not it is significant (Table III). Yield presents a positive correlation with rainfall from November to March. Meanwhile, in April and May, this correlation becomes indirect, only becoming significant in stations in the north and south extremities of the region. This shows that higher rainfall positive anomalies during the maturity-harvest period produce a negative impact on the final yield of the crop according to Pascale *et al.* (1983). The relation between the yield and the monthly mean maximum temperature is observed in a negative manner from January to April, while the correlation between the yield and the monthly mean minimum temperature stands out mainly in positive form in the first months (October and November) in the southern-most areas. The association between yield and the monthly mean thermal amplitude is negative and statistically significant in the areas located more to the north (Table III). These results are statistically significant, showing the sensitivity of the soybean crop to different extremes of climatic conditions (high and low temperatures, low humidity) in different crop cycle stages (da Mota, 1978).

The greatest significant regional association with the yield was shown by the seasonal precipitation that comprises the whole core and west region (Figure 5), thus accounting for up to 32% of the soybean yield variance. Even if the rainfall effect in the growth of the crop is not homogeneous (with different degrees of influence according to the phenologic stage), this result would indicate that seasonal rainfall in the core and western zones may be taken as a good initial yield indicator, in concurrence with the results obtained by other authors (Krepper *et al.*, 1998; Scian, 2002).

3.4. Crop-climate relationships

The joint effect of the climatic variables in the final yield of the crop was analysed using stepwise multiple

Table III. Correlations between the anomalies of the climatic variables and yield. (**) 95% and (*) 90% significance.

	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Marcos Juárez								
Rainfall	0.22	$\mathbf{0.36}^{*}$	$\mathbf{0.41}^{**}$	0.25	$\mathbf{0.41}^{**}$	0.23	−0.07	−0.22
Max T	0.03	−0.15	−0.07	$\mathbf{-0.50}^{**}$	$\mathbf{-0.36}^{*}$	−0.21	$\mathbf{-0.44}^{**}$	−0.12
Min T	0.32	0.32	$\mathbf{0.36}^{*}$	−0.06	0.09	0.15	−0.15	$\mathbf{-0.33}^{*}$
ΔT	−0.23	$\mathbf{-0.37}^{*}$	$\mathbf{-0.37}^{*}$	$\mathbf{-0.63}^{**}$	$\mathbf{-0.44}^{**}$	$\mathbf{-0.39}^{**}$	−0.21	0.27
Pilar								
Rainfall	$\mathbf{0.49}^{**}$	0.07	0.29	0.14	0.15	0.07	0.31	0.10
Max T	−0.23	$\mathbf{-0.36}^{*}$	−0.24	$\mathbf{-0.39}^{**}$	−0.32	−0.02	$\mathbf{-0.66}^{**}$	−0.25
Min T	0.20	0.00	0.01	−0.15	−0.01	0.18	−0.02	−0.15
ΔT	−0.32	$\mathbf{-0.50}^{**}$	−0.30	$\mathbf{-0.37}^{*}$	−0.31	−0.20	$\mathbf{-0.52}^{**}$	−0.09
Córdoba								
Rainfall	0.37	0.10	$\mathbf{0.55}^{**}$	0.25	−0.02	0.09	0.16	−0.27
Max T	0.12	−0.15	−0.31	−0.30	−0.29	−0.06	−0.20	0.26
Min T	0.07	−0.02	−0.14	−0.13	−0.02	0.18	0.03	0.08
ΔT	0.05	−0.17	−0.28	−0.32	$\mathbf{-0.35}^{*}$	−0.23	−0.06	0.19
Rosario								
Rainfall	0.30	$\mathbf{0.55}^{**}$	$\mathbf{0.55}^{**}$	0.28	0.09	$\mathbf{0.37}^{*}$	−0.15	−0.17
Max T	0.06	−0.33	−0.29	$\mathbf{-0.65}^{**}$	$\mathbf{-0.35}^{*}$	0.00	$\mathbf{-0.55}^{**}$	0.05
Min T	$\mathbf{0.49}^{**}$	0.28	0.20	−0.30	0.06	0.29	−0.24	−0.30
ΔT	$\mathbf{-0.39}^{*}$	$\mathbf{-0.45}^{**}$	$\mathbf{-0.46}^{**}$	$\mathbf{-0.49}^{**}$	$\mathbf{-0.33}^{*}$	$\mathbf{-0.34}^{*}$	−0.13	$\mathbf{0.42}^{**}$
Reconquista								
Rainfall	−0.12	−0.22	−0.06	−0.05	0.31	0.29	0.29	0.11
Max T	0.05	0.19	0.17	0.17	−0.08	−0.22	$\mathbf{-0.45}^{**}$	−0.21
Min T	−0.33	−0.21	−0.05	0.11	0.13	−0.13	−0.13	$\mathbf{-0.34}^{**}$
ΔT	0.32	0.28	0.25	0.08	−0.13	−0.14	−0.25	0.20
Ceres								
Rainfall	0.19	0.16	0.12	0.19	0.01	0.21	$\mathbf{-0.38}^{*}$	−0.26
Max T	−0.10	−0.01	0.01	−0.24	−0.21	−0.06	−0.18	−0.02
Min T	0.10	0.25	0.30	0.03	0.07	0.03	−0.24	−0.31
ΔT	−0.15	−0.11	−0.32	−0.22	−0.26	−0.10	0.07	$\mathbf{0.38}^{*}$
Junín								
Rainfall	0.25	$\mathbf{0.51}^{**}$	0.22	0.22	0.13	0.19	$\mathbf{-0.34}^{*}$	−0.27
Max T	0.30	0.15	−0.28	−0.19	−0.13	$\mathbf{-0.41}^{**}$	$\mathbf{-0.36}^{*}$	−0.01
Min T	$\mathbf{0.59}^{**}$	$\mathbf{0.51}^{**}$	0.20	−0.08	0.28	0.04	−0.08	−0.17
ΔT	−0.23	−0.24	$\mathbf{-0.40}^{**}$	−0.12	$\mathbf{-0.34}^{*}$	$\mathbf{-0.49}^{**}$	−0.20	0.18
9 de Julio								
Rainfall	0.32	0.18	−0.08	0.01	0.00	0.29	$\mathbf{-0.60}^{**}$	0.10
Max T	0.12	0.17	0.04	−0.16	0.08	$\mathbf{-0.58}^{**}$	−0.04	$\mathbf{-0.35}^{*}$
Min T	$\mathbf{0.41}^{**}$	$\mathbf{0.33}^{*}$	0.19	0.04	$\mathbf{0.41}^{**}$	−0.14	−0.32	−0.24
ΔT	−0.27	−0.09	−0.11	−0.26	−0.25	$\mathbf{-0.53}^{**}$	0.20	−0.02
Pehuajó								
Rainfall	0.04	−0.22	−0.10	$\mathbf{0.35}^{*}$	0.20	0.11	−0.16	0.09
Max T	−0.08	$\mathbf{0.57}^{**}$	0.30	−0.24	0.02	$\mathbf{-0.44}^{**}$	0.20	0.22
Min T	−0.02	0.20	0.23	−0.03	−0.05	$\mathbf{-0.61}^{**}$	−0.26	0.00
ΔT	−0.09	0.32	0.19	−0.18	0.09	0.17	$\mathbf{0.35}^{*}$	0.18
Paraná								
Rainfall	0.07	−0.04	$\mathbf{0.39}^{**}$	$\mathbf{0.51}^{**}$	−0.16	$\mathbf{0.60}^{**}$	−0.07	−0.20
Max T	−0.16	−0.13	−0.08	$\mathbf{-0.62}^{**}$	$\mathbf{-0.46}^{**}$	−0.25	−0.29	−0.13
Min T	−0.01	−0.18	0.21	−0.43	-0.47^{**}	−0.14	−0.24	−0.16
ΔT	−0.11	0.05	−0.27	−0.32	−0.08	−0.16	0.01	0.05
Gualeguaychú								
Rainfall	0.02	0.07	0.21	0.25	0.11	0.29	−0.28	$\mathbf{-0.39}^{**}$
Max T	−0.02	−0.21	−0.06	$\mathbf{-0.42}^{**}$	$\mathbf{-0.45}^{**}$	−0.09	−0.33	−0.23
Min T	0.06	0.07	0.26	−0.25	0.05	0.19	−0.07	−0.25
ΔT	−0.05	−0.27	$\mathbf{-0.34}^{*}$	−0.32	$\mathbf{-0.44}^{**}$	$\mathbf{-0.39}^{**}$	−0.05	0.10

Min T: Minimum Temperature, Max T: Maximum Temperature, ΔT: Thermal Range

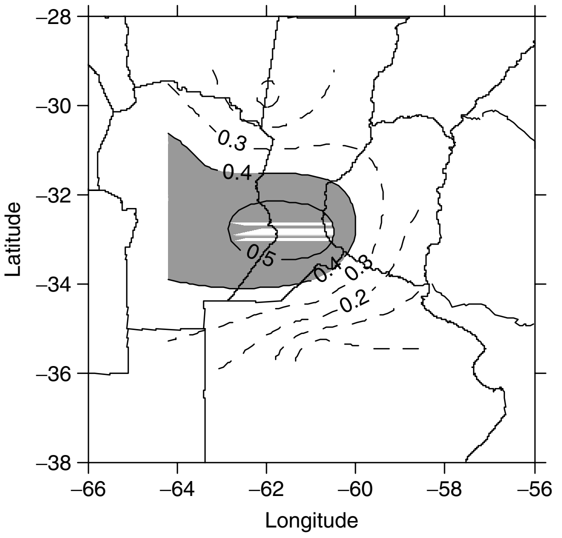

Figure 5. Correlation field between seasonal rainfall and soybean yield. Shaded area: Significant correlation at the 95% confidence level.

regression. The predictor variables proposed in each adjustment were selected according to the different months of the crop cycle stages, the variables mentioned in works by other authors (see Introduction), significant correlations between the yield and the climatic variables (analysed in Section 3.3).

Regression models are shown below for two stations located in the soybean core area, Rosario (Equation (1)), Marcos Juárez (Equation (2)), and one in the south, Junín (Equation (3)):

$$\Delta R_{\mathrm{ROS}} = -0.77 - 81.4\Delta T \underset{\mathrm{Jan}}{\max} + 2.1\Delta PP_{\mathrm{Nov}} - 97.4\Delta T \underset{\mathrm{Apr}}{\max} + 1.5\Delta PP_{\mathrm{Dec}} \quad (1)$$

$$\Delta R_{\mathrm{MJ}} = 0.41 - 148.7\Delta T \underset{\mathrm{Jan}}{\max} + 170.2\Delta T \underset{\mathrm{Dec}}{\min} + 184.6\Delta T \underset{\mathrm{Feb}}{\min} - 75.9\Delta T \underset{\mathrm{Dec}}{\max} \quad (2)$$

$$\Delta R_{\mathrm{JU}} = 0.37 + 167.7\Delta T \underset{\mathrm{Oct}}{\min} - 125.6\Delta T \underset{\mathrm{Mar}}{\max} + 1.36\Delta PP_{\mathrm{Mar}} + 1.78\Delta PP_{\mathrm{Nov}} - 72.4\Delta T \underset{\mathrm{Apr}}{\max} \quad (3)$$

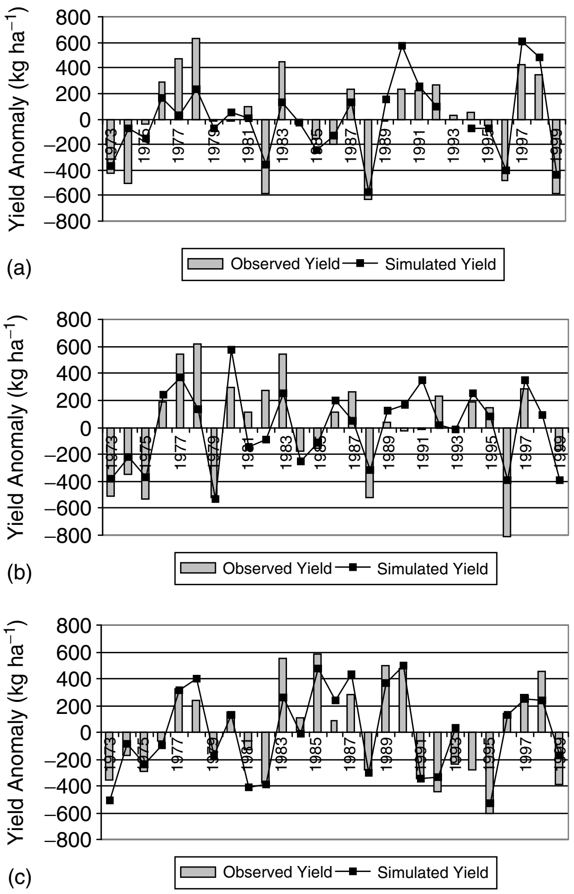

Figure 6. Observed and simulated yields anomalies of soybean during the period 1973–2000 for Rosario (a), Marcos Juárez (b), Junín (c).

Table IV. Statistics of the developed multilinear models. R^2: coefficient of determination; df: degrees of freedom; RMSE: root mean square error; SD: standard deviation.

Station	R	R^2	df	RMSE (kg ha^{-1})	SD (kg ha^{-1})
Rosario	0.83	0.68	4/21	195.4	354.6
Marcos Juárez	0.79	0.63	4/22	216.0	361.9
Junín	0.91	0.84	4/21	136.1	350.3

Where ΔR_{ROS}, ΔR_{MJ} and ΔR_{JU} are simulated anomalies of de-trended soybean yield in Rosario (Station 7), Marcos Juárez (Station 6) and Junín (Station 9) respectively; $\Delta T\,\max_{\mathrm{Dec}}$, $\Delta T\,\max_{\mathrm{Jan}}$, $\Delta T\,\max_{\mathrm{Mar}}$ and $\Delta T\,\max_{\mathrm{Apr}}$ are anomalies of maximum air temperature in December, January, March and April; $\Delta T\,\min_{\mathrm{Oct}}$, $\Delta T\,\min_{\mathrm{Dec}}$ and $\Delta T\,\min_{\mathrm{Feb}}$ are anomalies of minimum air temperature in October, December and February; and ΔPP_{Mar}, ΔPP_{Nov} and ΔPP_{Dec} are anomalies of rainfall in March, November and December. In Figure 6 the anomalies of the observed and estimated yield (Equations (1–3)) are shown. The variabilities of the observed and estimated yield anomalies resulting from these models are in phase with r^2 ranging from 0.63 to 0.84 (Table IV). In all the cases the adjustment was statistically significant at the 5% level. A comparison with the Root Mean Square Errors (RMSE) shows that the standard deviation values are higher than the RMSE, showing the utility of the prediction of annual soybean yield from climatic variables (Table IV).

In Rosario (Station 7), the yield is directly related to the rainfall during November and December and indirectly to the maximum temperatures in January and April. The maximum temperature in January is related in implicit form to the precipitation in January owing to the significant association between them. This result indicates the strong dependence of the yield on water availability, mainly determined by precipitation and temperature. The maximum temperature in April is related to the fact that it shows a significant association with the maximum temperature in February and the humidity in March. The negative relation between yield and the maximum temperature in January and February (through $\Delta T\ \max_{\mathrm{Apr}}$), appropriately represented in the equation, is the result of the sensitivity of the soybean plant to high temperatures (Palmer *et al.*, 1995).

In Marcos Juárez (Station 6), positive anomalies in the maximum temperature of December and January, produce a decrease in the final yield. This result again shows the sensitivity of soybean to high temperatures during the summer months. Both temperatures present significant correlations with the rainfall and thermal range of the respective months. In the same way, the relation between the yield and precipitation in February is represented by the minimum temperature of that month. Positive anomalies in the minimum temperature of December favour the soybean final yield in Marcos Juarez (Station 6) given these conditions are favourable for the growing process.

In Junín (Station 9), the November positive rainfall anomalies favour the final yield since these are associated with soil water availability (Hurtado *et al.*, 2001). November rainfall in this location is significantly correlated to the same variable in October. The water availability in the soil is reflected in the same way by the minimum temperature of October being significantly associated with the precipitation and thermal amplitude of the same month (pre-sowing). Negative anomalies in rainfall and positive ones in maximum temperature in March (warm and dry conditions) produce a decrease in the yield.

In all the cases, yield has a strong association with monthly extremes of temperature at important phenological moments of the crop, which, in turn, show a significant correlation with the monthly rainfall. Moreover, the predictor variables for each regression model depend on the location. The same result was observed in the other analysed locations not presented in this work.

4. Conclusions

The relationship between climatic factors and soybean yield fluctuation in the Argentine Pampas was analysed. To perform this analysis the monthly mean maximum and minimum temperatures, monthly mean thermal range, monthly rainfall and seasonal rainfall variability were examined in relation to soybean yield during the years 1973–2000. Moreover, the relationship between the soybean yield and these climate variables was quantified. The study region is the zone of larger production of this crop within the humid Argentine Pampas and contributes to a high percentage to the global trade of soybean and its by-products.

Even if annual rainfall and monthly mean maximum and minimum temperatures have significant positive trends in the last 50 years, in the period of study, only the minimum November temperature has shown this low-frequency behaviour in a spatially generalized manner. This suggests that there is a tendency to favourable climate conditions at the beginning of the soybean growing season. The yield presented a significant positive trend in 72% of the locations studied. This effect could be due to technological and/or climatic factors. According to this hypothesis and because the medium- and short-term variability are of interest in this work, the linear trend was filtered out in climatic and agronomic variables.

The de-trended soybean yield interannual variability is in phase with the seasonal rainfall interannual variability. This dependence does not appear in some extreme rainfall situations since years with negative as well as positive extreme anomalies of seasonal rainfall presented losses in the soybean crop, severe in some cases. An important consequence for the region, especially from the agricultural standpoint is that the occurrence of these extreme rainfall anomalies shows up in isolated years with a low probability of occurrence.

The analysis of the crop's spatial homogeneity, shown by the correlation fields, suggests that the whole region could be considered as composed of a series of sub-regions. The homogeneity of each sub-region is dependent on anthropogenic factors and/or the sub-region's natural characteristics, which define hydro-climatic macro systems.

The relationship between each climatic variable and the yield did not show a defined regional pattern. The simple correlations between these variables confirm in quantitative terms that air temperature and precipitation are the major climatic factors that determine the variability of crop production. However, the degree of association between each of these climatic variables and the yield is not significant, since they separately explain very little variance.

On the basis of these results, the joint effect of the climatic variables was studied for the soybean yield through a stepwise multiple regression model, supposing that the effects of other factors (for example soil, technology and pests) are contained in the residuals. In general, models explain a high percentage of the variability of the de-trended yield (more than 62% of yield variance). It can be inferred that the combination of monthly extreme temperatures and precipitation has a prevailing role in the yield definition. It is interesting to note that rainfall does not always appear explicitly in the statistical models. Its effect, however, is represented implicitly through the high partial correlation with temperatures.

At a sub-region scale, these analyses showed that the soybean crop is sensitive to climatic variability. In general, a higher maximum temperature during summer months and rainfall excesses in the maturity-harvest period normally result in a lower yield, while a higher minimum temperature during the growing season can increase soybean yield. The crop's negative dependence on atmospheric humidity is shown significantly during summer months at the stations located in the north. The most significant spatial coherence with the yield was shown by seasonal precipitation, which can be considered a proper yield indicator.

The description of the relation between climate and soybeans in the humid Argentine Pampas is proposed as an initial step in the chain of studies necessary to evaluate the impact of climate variability and change on production. This paper sets the bibliographic basis of statistical modelling of the climate-related crop, and quantifies objectively the spatial and temporal variability of soybean yield that can be allocated to climate variability in the region. The outcome of this study can be used by researchers to assess current climatic fluctuations and variability and the expected climate change for the twenty-first century.

The heterogeneity of the yield suggests that every state or group of producers should have their own strategy. In general, these strategies depend on several factors (rates, policies and technological availability) and also on climate and weather. The relationship between climatic variables and yield should be analysed in a decision process. Thus, it is possible to determine which parameters need to be monitored and therefore to minimize the investment required for the expected yield. The results obtained in this study can also be useful in the ongoing Argentinian national assessment of the potential consequences of climate variability and change on agricultural production, land-use and natural resources. These approaches may be very useful to a broad range of decision makers in the agricultural sector.

Acknowledgements

The authors wish to thank the two anonymous reviewers for their very useful comments that improved this paper. This work was supported by the following grants from Universidad de Buenos Aires X135 and X234, CONICET PIP 5139 and CLARIS Project (European Commission Project 001454).

References

Abbott PF. 1986. *Guidelines on the Quality Control of Surface Climatological Data*. World Meteorological Organization: Geneva; WCP-85, WMO/TD-No. 111; 65.

Alexandrov VA, Hoogenboom G. 2001. Climate variation and crop production in Georgia, USA, during the twentieth century. *Climate Research* **17**: 33–43.

Barrucand M, Rusticucci M. 2001. Climatología de temperaturas extremas en la Argentina. Variabilidad temporal y regional. *Meteorológica* **26**: 85–102.

Bolsa de Cereales. 1979. *Revista Institucional. Número Estadístico*, 328.

Boullón D. 2002. *Evaluación del riesgo en la Región Pampeana*. SAGPyA (Secretaría de Agricultura, Ganadería, Pesca y Alimentación de la Nación) Oficina de Riesgo Agropecuario. http://www.sagpya.mecon.gov.ar/10/informes$_{01}$.htm.

Brooks CEP, Carruthers N. 1953. *Handbook of Statistical Methods in Meteorology*. Her Majesty's Stationery Office: London.

Calviño PA, Sadras VO. 1999. Interanual variation in soybean yield: interaction among rainfall, soil depth and crop management. *Field Crops Research* **63**: 237–246.

Castañeda E, Barros V. 1994. Las tendencias de la precipitación en el Cono Sur de América al este de los Andes. *Meteorológica* **19**: 23–32.

da Mota FS. 1978. *WMO Technical Note N°160*. Soya Bean and Weather, 64.

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

Easterling D, Horton B, Jones P, Petterson T, Karl T, Parker D, Salinger MJ, Rasuvayev V, Plummer N, Jamason P, Folland C. 1997. Maximum and minimum temperature trends for the globe. *Science* **277**: 364–367.

Garcia P, Offutt SE, Pinar M, Changnon SA. 1987. Corn yield behavior: effects of technological advance and weather conditions. *Journal of Climate and Applied Meteorology* **26**: 1092–1102.

Giorda LM, Baigorri H. 1997. *El Cultivo de la Soja en Argentina*. INTA, SAGPyA: Buenos Aires, Argentina; 448.

Hall AJ, Rebella CM, Ghersa CM, Culto JP. 1992. In *Field-crop Systems of the Pampas. Ecosystems of the World: Field-crop Ecosystems*, Pearson CJ (ed.). Elsevier: Exeter, UK; 413–449.

Höel PG. 1964. *Introduction to Mathematical Statistics*, 3rd edn. John Wiley and Sons: New York; 428.

Hoffmann JA. 1975. *Atlas Climático de América del Sur*. WMO: Hungary.

Hoffmann JA, Núñez S, Gómez A. 1987. Fluctuaciones de la precipitación en la Argentina, en lo que va del siglo. *Preprints del II Congreso Interamericano de Meteorología. V Congreso Argentino de Meteorología*. Centro Argentino de Meteorólogos: Buenos Aires; 12.1.1–12.1.5.

Huff FA, Neill JC. 1982. Effects of natural climatic fluctuations on the temporal and spatial variation in crop yields. *Journal of Applied Meteorology* **21**: 540–550.

Hurtado R, Fernández Long ME, Spescha L, Murphy G. 2001. *Distribución espacial de los rendimientos extremos de soja en la Región Pampeana y condiciones climáticas que los acompañan*. Actas de XII Congresso Brasileiro de Agrometeorologia y III Reuniao Latino-Americana de Agrometeorologia, 3 a 6 de julho de 2001, Fortaleza, Brasil.

Krepper CM, Paoloni JD, Sequeira ME. 1998. Variabilidad de baja frecuencia de las precipitaciones y su influencia en el desarrollo del cultivo de soja en Argentina. *Preprints de la III Reunión Nacional de Oleaginosos*. Asociación Argentina de Oleaginosos: Bahía Blanca; 161–162.

Magrin GO, Grondona MO, Travasso MI, Boullón DR, Rodríguez GR, Messina CD. 1998. *Impacto del fenómeno "El Niño" sobre la producción de cultivos en la región pampeana*. Instituto de Clima y Agua. INTA Castelar: Buenos Aires, Argentina; 16.

Minetti JL, Lamelas CM. 1995. Respuesta regional de la soja en Tucumán a la variabilidad climática. *Revista Industrial y Agrícola de Tucumán* **72**: 63–68.

Palmer J, Dunphy EJ, Reese P. 1995. *Managing Drought-stressed Soybeans in the Southeast, North Carolina Cooperative Extension Service*. North Carolina State University: Raleigh, NorthCarlino, USA.

Panofsky H, Brier G. 1965. Some applications of Statistics to Meteorology. *College of Mineral Industries*. The Pennsylvania State University: Pennsylvania, USA; 224.

Pascale AJ, Escales EA. 1971. Requerimientos bioclimáticos de los cultivares de soja. I: subperíodo siembra-floración. *Revista de la Facultad de Agronomía y Veterinaria* **19**: 15–24.

Pascale AJ, Escales EA. 1975. Requerimientos bioclimáticos de los cultivares de soja. II: subperíodo floración-maduración. *Revista de Información sobre Investigación y Desarrollo Agropecuario*. **31**: 53–63.

Pascale AJ, Murphy GM. 1975. Condiciones agroclimáticas para la maduración de la soja en la subregión oriental. *Revista de Información sobre Investigación y Desarrollo Agropecuario*. **31**: 40–53.

Pascale AJ, Rodríguez RO. 1977. Aptitud mesoagroclimática/hídrica de la subregión oriental para el cultivo de la soja. *Preprints de la V Reunión Técnica Nacional de Soja*. Asociación Argentina de Oleaginosos: Miramar; 414–424.

Pascale AJ, Miaczynski CR, Rodríguez RO, Trigo ER. 1983. Aptitud ecológica de la región oriental argentina para el cultivo de la soja. *Preprints de la VIII Reunión Técnica Nacional de Soja*. Asociación Argentina de Oleaginosos: Tucumán; B1–B-19.

Penalba OC, Vargas WM. 1996. Climatology of monthly and annual rainfall in Buenos Aires, Argentina. *Meteorological Applications* **3**: 275–282.

Penalba OC, Vargas WM. 2001. Propiedades de precipitaciones extremas en zonas agropecuarias argentinas. *Meteorológica* **26**: 39–55.

Penalba OC, Vargas WM. 2004. Interdecadal and Interannual variations of annual and extreme precipitation over central-northeastern Argentina. Changes in the extreme precipitation seasonal cycle. *International Journal of Climatology* **24**(12): 1565–1580.

Podestá GP, Messina CD, Grondona MO, Magrín GO. 1999. Associations between Grain Crop Yields in Central-Eastern Argentina and El Niño-Southern Oscillation. *Journal of Applied Meteorology* **38**: 1488–1498.

Ravelo AC, Pascale AJ, Planchuelo AM. 1983. Modelo para estimar los estados fenológicos de la soja. *Preprints de la VIII Reunión Técnica Nacional de Soja*. Asociación Argentina de Oleaginosos: Tucumán; B65–B71.

Rusticucci MM, Penalba OC. 2000. Interdecadal changes in the precipitation seasonal cycle over Southern South America and their relationship with surface temperature. *Climate Research* **16**: 1–15.

SAGPyA. 2000. *Estimaciones Agrícolas*. Ministerio de Economía y Obras y Servicios Públicos, Secretaría de Agricultura, Ganadería, Pesca y Alimentación de la Nación Argentina.

Schwerdtfeger W. 1976. *Climates of Central and South America. World Survey of Climatology*, Vol. 12. Elsevier Scientific Publishing Company: The Netherlands; 532.

Scian B. 2002. Variabilidad de las condiciones hídricas en la region semiárida pampeana, Argentina. *Geoacta* **27**: 30–52.

Wilks DF. 1995. *Statistical Methods in the Atmospheric Sciences*. Academic Press: San Diego, CA, USA; 467.

Yao AYM. 1981. *General Climatology*, World Survey of Climatology, Vol. 3. Elsevier Scientific Publishing Company: The Netherlands; 402.

DOI: 10.1002/met