ELSEVIER
Journal of Arid Environments 61 (2005) 669–679
Journal of Arid Environments
www.elsevier.com/locate/jnlabr/yjare

Environmental variables related to wheat yields in the semiarid pampa region of Argentina

B.V. Scian*, M.E. Bouza

Dpto. de Agronomía, Universidad Nacional del Sur, 8000 Bahía Blanca, Argentina

Received 11 April 2003; received in revised form 31 August 2004; accepted 14 September 2004

Abstract

The current paper analyses various environmental parameters in relation to wheat yields in Bordenave, Province of Buenos Aires, Argentina. The variables used are: precipitation (ppt), maximum (Tx) and minimum temperature (Tmi) as well as those obtained by applying the Palmer model. Decadic and phenological scales are used for data corresponding to the period 1977–2000. The stepwise method is used to obtain a multi-variate equation to calculate yield taking into account environmental variables only. For a five variates model the coefficient of determination, R^2 equals 95.79% and the standard error of estimation is 129.0 kg ha^{-1}. In the sample yields, the incidence of total variability for thermal variables is 42.7% and for hydrological variables, 53%. The value and sign of the correlation coefficients were analysed throughout the cultivation cycle. The α coefficient is mainly responsible for yield variance during tillering and stem elongation. There is good correlation with the values of Palmer's Drought Severity Index (PDSI) for the flowering and grain filling stages.

Keywords: Environmental variables; Wheat yield; Correlations; Semiarid pampa region

---

*Corresponding author. Tel.: +54 291 4595102x4360; fax: +54 291 4595127.
E-mail address: mtscian@criba.edu.ar (B.V. Scian).

0140-1963/$ - see front matter
doi:10.1016/j.jaridenv.2004.09.018

1. Introduction

The semiarid pampa region is characterized by the variability of its annual precipitation, average annual precipitation varying between 500 and 700 mm, depending on the latitude, and by water deficit throughout the whole year. Various authors point out that the winter cereals in the region suffer from severely limited water availability mainly in the month of November. According to Paoloni and Vazquez (1984) the highest correlations between monthly rainfall and annual wheat yields occur in November and during fallow (March and April). Of particular importance is the initial water content in the soil at the time of sowing and water deficiency during anthesis (Travasso et al., 1994). Average wheat yields in the semiarid region are 2000 kg ha^{-1}, with maxima of 3500–3600 kg ha^{-1} in cases of early and plentiful rainfall occurring between the end of November and December, no late frosts and moderate temperatures (Gallo Candolo, 2001). Regression-type models have been used since the beginning of the 20th century to assess the effect of weather on yield, and they have also been criticized for their empirical nature, which restricts their application to the range of values for which they were developed. Nevertheless, such models are still adequate for regions where simulation or deterministic growth models are not available, since apart from other factors, it is not possible to correctly determine the appropriate model parameters. Furthermore there is renewed interest in this type of model as a tool for calculating the impact of climatic fluctuations on crop production over wide-spread areas. Early investigations using monthly moisture anomaly index (Z) to estimate wheat yield performed on Bordenave data showed that the selected Z variables corresponded to the months of October and November (Scian and Donnari, 1995).

The aim of the current work is to determine which environmental variables have the greatest incidence on wheat yield and then use these to make predictions on a decadic scale in a multi-variate statistical model for the Bordenave region in the Province of Buenos Aires.

2. Materials and methods

2.1. Study area

The data used correspond to the Bordenave Agricultural Experimental Station (Latitude: 37°51′S, Longitude: 63°01′W, height: 212 m) located in the semiarid west border of the Province of Buenos Aires. The soil is Haplustol Entic, of fine sandy loam typical of the region, with medium to low fertility and susceptible to wind erosion. Soil depth—0.8 to 1.2 m—is determined by the calcareous stratum.

2.2. Agro-meteorological variables

Of the yield data available from a pilot experiment with long-cycle crops (Prointa Pigué, Cochicó Inta, Calden Inta) it was decided to use only those of a control experiment in which neither fertilizers nor crop management were applied (source: Ing. Venanzi, 2000). The meteorological data used are: daily precipitation values (ppt), minimum (Tmi) and maximum (Tx) temperatures, relative humidity, wind velocity at 2 m, and relative heliophany corresponding to the period 1977–2000 (source: Ing. Campi, 2001). No yield data are available for the 1993 and 1994 harvests. The Palmer model (Palmer, 1965) was applied on a decadic scale, using an available water capacity of 150 mm. Potential evapotranspiration (PET) was calculated according to Penman-FAO. Palmer's moisture anomaly (Z) and Drought Severity (PDSI) indices were used, as well as soil moisture available (Smt) and climatic coefficient α of Palmer's water balance, which is the ratio of real evapotranspiration (rET)/PET.

2.3. *Phenological stages of wheat*

Sowing dates in the semiarid pampa region vary between June and the first-half of July, depending on the degree of probability of late frosts. For the purposes of the present study an average sowing date was determined coinciding with the third decade of June. Taking into account wheat growth patterns and adjusting the periods to decadic agroclimatic variables, the following stages were determined: (1) sowing, (2) emergence, (3) tillering, (4) stem elongation, (5) heading, (6) anthesis, (7) grain filling, and (8) maturity and harvesting. Table 1 shows the decadic period corresponding to each stage (S1–S8). The ten-day periods are identified by numbers 1, 2 and 3, indicating the first, second and third ten-day period of the month. The period from the first decade of March to the second decade of June corresponding to fallow was also taken into account.

2.4. *Statistical analysis*

The (r) correlations were calculated with the standard formulas of Pearson and a significance level set at 5%.

The multiple regression model for calculating yield gives the following equation:

$$\star = B_0 + \Sigma B_i X_i$$

where $\star$ is the estimated yield, B_0 the constant, B_i the coefficient of independent variables, X_i the decadic environmental variables.

Table 1
Phenological wheat stages (S) as mentioned in text (S1–S8) for Bordenave, average starting dates and decade

	Sowing (S1)	Emergence (S2)	Tillering (S3)	Stem elong (S4)	Heading (S5)	Anthesis (S6)	Grain filling (S7)	Maturity and harvest (S8)
START	Jun 27	Jul 18	Set 1	Oct 1	Oct 30	Nov 8/ 9	Nov 12	Dec 8
DECADE	jn3	jl2-ag3	se1-set 3	oc1-oc2	Oc3	no1	no2-no3	di1-di2

For the multi-variate regression analysis the stepwise method with an *F* value of 4 for the entry of variables was used, and a minimum tolerance between 0.01 and 0.20, depending on the need to restrict the entry of variables (Draper and Smith, 1981). The number of variables to be included in the model was fixed at 5, in order to have enough degrees of freedom to perform the regression. Management, genetic and technological variables were not taken into account. The first 16 years were used to obtain the model and the predictive capacity of the model was then tested on records from recent years that were separated for control purposes, using RMSE, which is more sensitive to large errors than absolute error (MAE).

3. Results and discussion

3.1. Relationship between variables

On the basis of the meteorological data for the 24 year-period covered, the average maximum and minimum temperatures (Tx and Tmi) and accumulated ppt for each of the 36 decades of the year were calculated (see Fig. 1). During crop development (June–December) both temperature and rainfall increase. Mean minimum temperature do not fall under 0 °C neither mean maximum overpass 30 °C. It is observed the high inter-decadic variability that is one of the causes of the variability in crop yield. The incidence of hydrological conditions on wheat can also be analysed as a function of the α coefficient, which represents the ratio between water availability and the plant's water requirements and is significant practically throughout the entire growth cycle. This latter parameter presents regional values that always remain $\geqslant 0.50$ (figure not shown) corresponding to the minimum water requirements for the establishment of a crop (Cócheme and Franquin, 1967).

The variability of wheat yields in Bordenave is strongly related to the amount of precipitation. Fig. 2 shows the percentage of accumulated rainfall during the growth

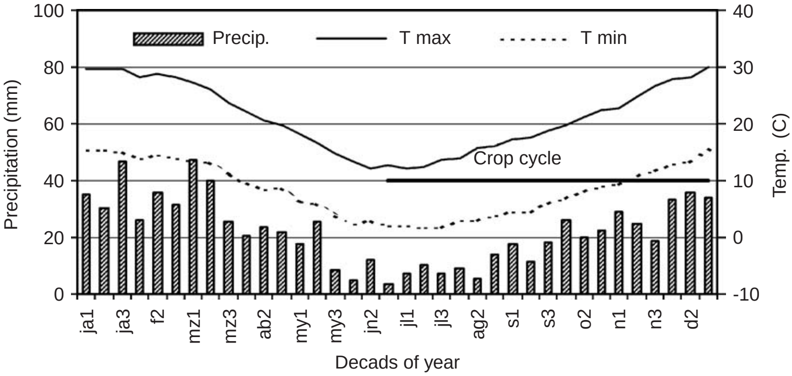

Fig. 1. Mean maximum and minimum temperature (C) and accumulated precipitation (mm) for the 36 decades of the year, Bordenave (1977–2000).

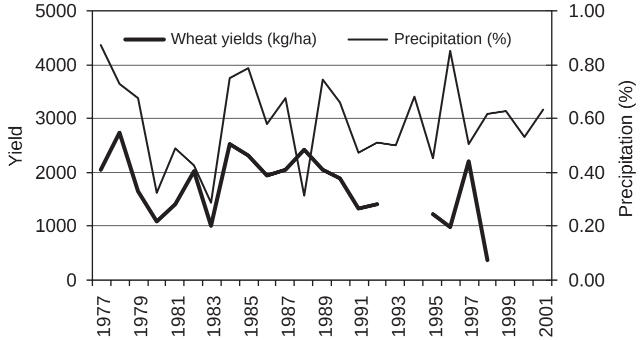

Fig. 2. Accumulated precipitation during wheat crop cycle as a percentage of annual total precipitation and grain yields (kg ha^{-1}) at Bordenave.

cycle of the crops with respect to total annual precipitation (second *y*-axis) together with the yields for the studied crop seasons. In general, a clear relationship can be observed between rainfall and yield. Despite the fact that 1988 stands out as a year with a low percentage of accumulated rainfall during the crop cycle, yields were higher than average. This behavior can be partly explained by the good pre-sowing storage of water in the soil (ppt (March) = 377 mm). 1996 was another exceptional year in which yield was contrary to expectations: the high percentage of rainfall was due mainly to the 372 mm which fell during the first decade of December, causing early sprouting of the grain and diminished yield.

The correlations for variables ppt, Tmi, Tx, α, Smt, Z index and the PDSI were calculated for the wheat yields, on a decadic scale and by phenological stage. The results are presented in Tables 2 and 3. One of the main factors impacting on the wheat crop cycle is the temperature—both maximum and minimum (Miralles and Slafer, 2001). A positive correlation with average Tmi during the vegetative period is observed, which is in agreement with the phenological concept that low temperatures during this stage favor the rate of onset of primordial leaves. The impact of the Tmi variable during the following growth stages is not statistically significant. It is not until flowering and up to ten days thereafter that low temperatures begin to negatively affect the crop, since it affects the final number of fertile flowers, thus reducing grain formation capacity.

The maximum temperatures during the stalk elongation and heading stages (month of October) also show a negative correlation with yield (Fig. 3), confirming that high temperatures during these stages are detrimental to wheat growth. Similarly, the occurrence of high temperatures at the commencement of the grain filling stage leads to the accumulation of dry matter in the grain, thus diminishing the overall duration of this stage and giving rise to the low final weight of the grain.

An analysis of the Smt, α, Z and PDSI variables shows that all have a significant effect on yield at one crop stage or another, and especially so at the vegetative stage (see Table 2). Fig. 4 shows the correlation between yield and each one of the variables. From sowing (June) to November there is a gradual increase in all

Table 2
Correlation coefficient (*r*) between wheat yields and environmental variables

r	ppt	Tmi	Tx	α	Smt	*Z*	PDSI
jn3	0.27	**0.57**	0.19	−0.14	−0.14	0.11	−0.14
jl1	−0.01	−0.03	−0.20	−0.07	−0.09	−0.15	−0.16
jl2	**0.51**	***0.60***	−0.01	0.21	0.06	**0.46**	0.15
jl3	0.33	0.26	−0.22	0.05	0.16	0.33	0.27
ag1	−0.12	0.16	−0.11	0.00	0.16	0.07	0.24
ag2	0.33	−0.13	−0.38	0.15	0.22	0.33	0.30
ag3	−0.03	**−0.57**	−0.21	0.33	0.14	−0.03	0.18
se1	0.08	0.08	−0.11	0.27	0.17	0.06	0.15
se2	−0.01	0.04	−0.07	0.18	0.19	0.02	0.13
se3	0.33	0.36	−0.07	**0.40**	**0.40**	***0.61***	0.38
oc1	0.27	0.08	***−0.51***	***0.59***	**0.45**	0.37	**0.52**
oc2	**0.48**	−0.13	***−0.60***	***0.72***	**0.51**	**0.54**	**0.61**
oc3	0.25	−0.07	−0.30	**0.53**	**0.51**	0.32	***0.60***
no1	**0.45**	−0.22	−0.22	**0.43**	**0.53**	**0.45**	***0.60***
no2	−0.18	−0.05	−0.37	0.24	**0.58**	**0.42**	***0.63***
no3	0.37	−0.09	−0.08	**0.43**	**0.61**	**0.54**	***0.69***
di1	−0.25	**−0.50**	−0.02	0.22	0.11	−0.30	0.05
di2	−0.21	−0.16	−0.15	−0.17	0.10	−0.05	0.02
di3	0.26	−0.24	0.00	−0.27	0.26	0.22	0.15

(Precipitation: ppt, Minimum Temperature: Tmi, Maximum Temperature: Tx, climatic coefficient: α, Soil moisture: Smt, Palmer's moisture anomaly: *Z*, and Drought Severity Index: PDSI) using decadal scale. Boldface values indicate significance level of 5%, italics for largest correlation in decade.

Table 3
Correlation coefficient for wheat yields and environmental variables as in Table 2, at phenological scale (5% statistical significant level in boldface)

r	ppt	Tmi	Tx	α	Smt	*Z*	PDSI
Sowing jn3–jl1	0.06	0.36	0.14	−0.12	−0.12	−0.07	−0.14
Emergence jl2–ag3	**0.37**	0.22	−0.30	0.16	0.15	0.31	0.06
Tillering se1–se3	0.27	0.23	−0.28	0.35	0.26	0.34	0.12
Stem elong. oc1–oc2	**0.54**	−0.04	**−0.55**	**0.70**	**0.48**	**0.59**	0.37
Heading oc3	0.25	−0.07	−0.26	**0.45**	**0.51**	0.32	**0.44**
Anthesis no1	**0.45**	−0.23	−0.31	0.36	**0.53**	**0.45**	**0.47**
Grain filling no2–no3	0.04	−0.09	−0.25	**0.60**	**0.60**	**0.61**	**0.53**
Maturity di1–di2	−0.29	**−0.41**	−0.09	0.07	0.11	−0.26	0.22
Harvest di3	0.26	−0.25	−0.02	−0.23	0.26	0.23	0.26

correlations. The highest statistically significant values are for the *Z* index in jl2 (*Z* jl2) and for the α variable in se3, oc1 and oc2 (α(se3), α (oc1) and α (oc2). Nevertheless, the other variables (Smt and PDSI) also correlate well from no1 to no3. The use of decadic scale instead of monthly or phenological, enables yield to be correlated with three times the amount of data, making it possible to determine in

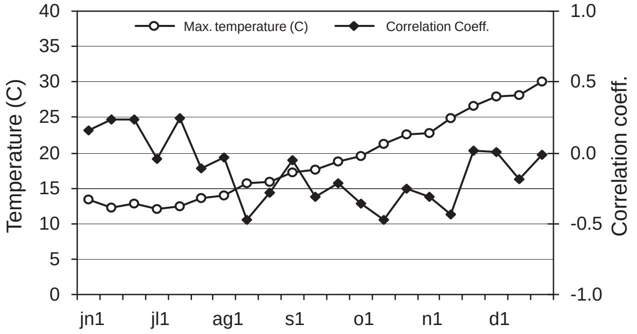

Fig. 3. Crop cycle mean maximum temperature (C) and yield correlation coefficients, Bordenave.

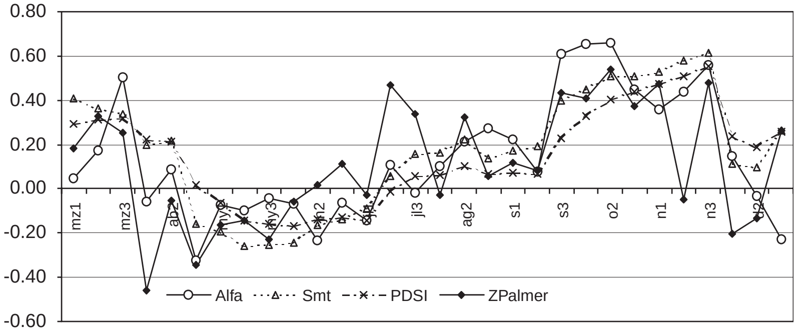

Fig. 4. Correlation coefficients between crop yields and environmental variables α, Smt (mm), PDSI and *Z*.

which fraction of the crop stage the relationship with environmental variables is higher.

Table 2 shows that the PDSI correlations reach the highest values from the end of October (oc3) to the end of November (no3), reflecting the importance during this period of water availability in the soil by means of a drought index. According to some authors the theoretical water requirement at this stage cannot be less than 100 mm (Travasso et al., 1994). Wheat yield deviations from the average recorded values (1808 kg ha^{-1}) were calculated and represented together with the progress of November PSDI in Fig. 5. The PDSI was designed as a meteorological index to measure soil water conditions after the computation of a simple soil water balance and highly autocorrelated with the previous moisture conditions. The correlation value between this index and yields proved to be 0.65. Similarly, in two localities at a distance of less than 50 km from Bordenave (17 de Agosto and Puan), very good correlations were achieved between estimated wheat yield and the PDSI for November, with correlation coefficients of 0.71 and 0.65, respectively.

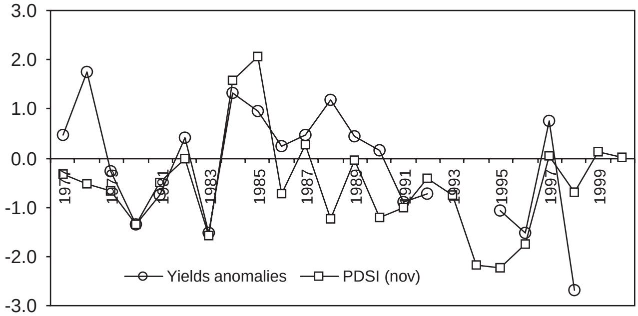

Fig. 5. Crop yield anomalies and November PDSI.

3.2. Multi-variate statistical model

Variables shown in Table 2 were analysed previously to their use in the definition of the statistic model. The cross correlation between variables for the whole crop cycle showed large values for the pairs: Z vs. PDSI (0.67) and Smt vs. α(0.67). Decision was taken to discard Smt and PDSI. Decades range from jn3 to nov3. The stepwise regression is the iterative application of forward selection and backward elimination steps. The idea is not so much to specify accurately or estimate a complete model, as it is to perform a selection of the factors that contribute significantly to variation in the response.

When the stepwise method was applied the results show the selected variables to be α (oc1), Tmi (ag3), ppt (no1), Tx (oc2) and Z (no2) with a coefficient of multiple determination, R^2, equal to 95.79% and the standard error of estimation of $128.76\,\mathrm{kg\,ha^{-1}}$. The number of regressors, p, in the model was sustained when another screening approach was applied, the Mallows' Cp statistic (Von Storch and Zweirs, 1999). A model that fits well will have a Cp close to p factors, that in our case amounts to 6. Table 4 shows the coefficients of the variables of the regression, B_0 to B_5, the standard error for these and the variation in R^2 for each step of the model after introducing a new variable. The use of cross-validation method allows to obtain a good estimate of future model performance (Von Storch and Zweirs, 1999). The skill of the yield model was calculated using the data of the years 1995, 1996 and 1997, and a RMSE of $316\,\mathrm{kg\,ha^{-1}}$ was obtained.

Although from the climatic point of view the sampling years were highly variable, owing to drought, frost and rainfall at inappropriate times for crop development, an equation was obtained using only environmental variables with an error of $300\,\mathrm{kg\,ha^{-1}}$ in yield assessment. This simple statistical model brings out the incidence of hydrological and thermal variables on yield. Fig. 6 shows the actual yields recorded and those estimated by the regression equation and Fig. 7 shows the residuals of the latter. The values indicate greater error for above average yields.

In general, the main factor explaining variability in crop yield is technological change, including the increased use of fertilizers, improved management practices

Table 4
Summary of Stepwise regression: Linear regression coefficients (B_i), Standard error of estimate, multiple determination coefficient R^2, variation of R^2, F to enter/rem and p-level

	B_i (kg ha^{-1})	St. Error (kg ha^{-1})	Multiple R^2	R^2 variation	F-to entr/rem	p-level
Indep. Coef.	3367.11	654.25				
α(oc1)	941.10	172.92	0.4377	0.4377	10.89886	0.0080
Tmi (ag3)	−133.61	17.30	0.8128	0.3751	26.05584	0.0005
ppt (no1)	4.36	1.46	0.8797	0.0669	6.67461	0.0272
Tx (oc2)	−91.72	25.08	0.9316	0.0519	8.34423	0.0161
Z(no2)	3.07	1.23	0.9579	0.0263	6.26291	0.0313

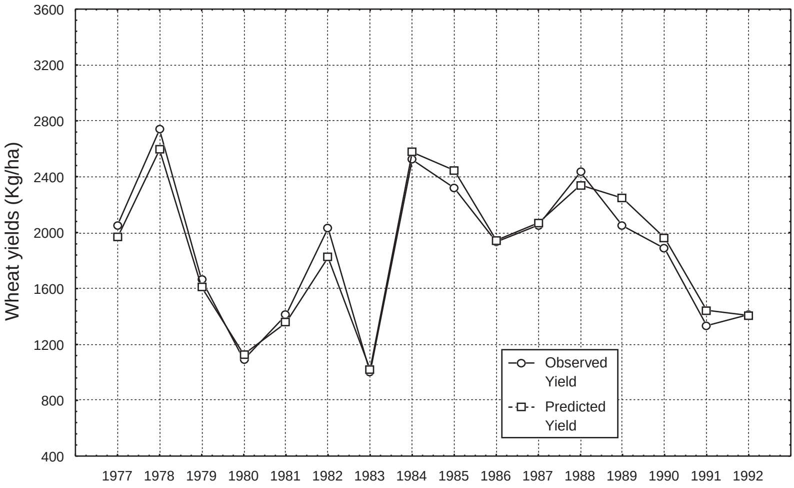

Fig. 6. Observed and predicted wheat yields, regression model.

and disease control, improvements in the genetic quality of the seeds, as well as other human interventions aimed at increasing yield. The second most important factor influencing variations in yield is meteorological variability within and between seasons (Baier, 1977). The yields observed in Bordenave (see Fig. 6) do not show a tendency to increase over the study period as a consequence of technological innovations or of climatic changes affecting hydrological and/or thermal conditions.

In other words, the variability in these yields is a consequence of variations in natural climatic conditions and it is interesting to analyse the participation of each variable in total variability. Analyzing R^2 (Table 4) shows that the variables related to precipitation are responsible for 53.1% of the climatic variability, those related to temperature (Tx and Tmi) are together responsible for 42.7%, and the rest can be attributed to residual errors. The 300 kg ha^{-1} RMSE obtained represents only the case of variability of natural elements on yield. Series of yields under different

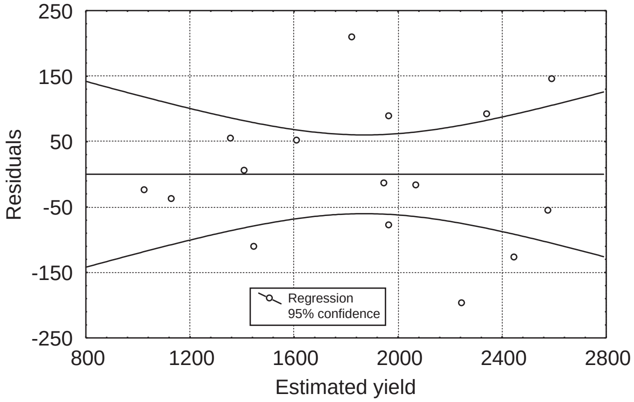

Fig. 7. Residuals as a function of estimated wheat yields.

treatments should therefore be analysed to establish the contribution of technological improvements and their incidence on error in yield assessment.

A brief comment about each one of the variables included in the model follows in order to help to relate their importance in crop yields. The inclusion of the coefficient α(oc1) is widely justified because in our study region this coefficient is greater than the critical value all year long, assuring the fulfillment of wheat hydric requirements. Besides, during the first decade of October stem elongation stage is in process and a demand in nutrients is increased. By the third decade of August the crop is at the end of its emergence phase so low temperatures, Tmi(ag3), are directly related to vernalization and thus promoting tillering. Precipitation regimen in the region presents a maximum in spring so that ppt(no1) is highly probable to get good rainfall conditions (except some ENSO years) during anthesis. Wheat crops during this stage are quite sensible to water deficit and thermal stress. If rainfall deficit conditions are present, water requirements should be provided by soil moisture availability. The statistical model includes the Tx(oc2) by the end of stem-elongation stage, while the greatest incidence should be assigned to the grain filling stage. As we can see in Fig. 3, there are two decades (oc2 and no2) in which high negative correlation exists between maximum temperature and crop yield. Finally, the other variable included in the model is Palmer's moisture anomaly during the second decade of November, Z(no2). This period corresponds to grain filling stage when the potential size of the grains is defined. External environmental factors as severe water stress or sudden heat shock may drastically terminate grain filling.

The results of the application of complex yield simulation models such as CERES–wheat for the pampa region indicate that the difference between real and estimated yields do not exceed one standard deviation (660 kg ha^{-1}), (Magrin et al., 1991). The application of a simple regression model as presented here could therefore be of practical value in predictions and a tool for assessing the impact of significant changes in environmental variables.

Acknowledgement

The collaboration of the student Romina Zuain is gratefully acknowledged. This work was partially financed by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) through grant PID-454/98.

References

Baier, W., 1977. Crop-Weather models and their use in yield assessments, Technical note No 151. World Meteorological Organization, Geneva, Switzerland, 48pp.

Cocheme, J., Franquin, P., 1967. An agroclimatological survey of a semiarid area in Africa South of the Sahara, Technical Note 86, WMO Nro. 210, 136pp.

Draper, N.R., Smith, H., 1981. Applied Regression Analysis, second ed. Wiley, New York 709pp.

Gallo Candolo, E., 2001. Zona Semiárida. In: Cuaderno de actualización 63, Consorcio Regional de Experimentación Agrícola (CREA), Buenos Aires, pp. 139–142.

Magrin, G., Díaz, R., Rebella, C., del Santo, C., Rodríguez, R., 1991. Simulación del crecimiento y desarrollo de trigo en Argentina y la necesidad de información meteorológica de entrada. Anales, Congremet VI, Buenos Aires, Argentina, pp. 49–50.

Miralles, D.J., Slafer, G.F., 2001. Desarrollo, crecimiento y determinación de las componentes del rendimiento de trigo. In: Cuaderno de actualización Nro. 63, Consorcio Regional de Experimentación Agrícola (CREA), Buenos Aires, pp. 8–17.

Palmer, W.C., 1965. Meteorological drought. U.S Dept. Commerce, Weather Bureau, Res. Paper 45, 58pp.

Paoloni, J.D., Vazquez, R., 1984. Necesidades teóricas de agua de los cereales de invierno y probabilidad de ocurrencia de las precipitaciones como base para el balance hídrico. Anales de Edafología y agrobiología, vol. XLIII, 11–12, Madrid, pp. 1545–1556.

Scian, B., Donnari, M., 1995. El índice de anomalía de humedad z y su relación con el rendimiento de trigo en Bordenave, Buenos Aires, Argentina. Agriscientia XII, 47–51.

Travasso, M.I., Dellecolle, R., Baldy, Ch., 1994. Principales factores agroclimáticos determinantes del rendimiento del trigo en el sur bonaerense. RIA 25 (2), 33–40.

Von Storch, H., Zweirs, F., 1999. Statistical Analysis in Climate Research. Cambridge University Press, United Kingdom 484pp.