

Forecast of frost days based on monthly temperatures

M. T. Castellanos^{1*}, A. M. Tarquis^{1,2}, M. C. Morató¹ and A. Saa²

¹ Dpto. de Matemática Aplicada a la Ingeniería Agronómica. ETSI Agrónomos. UPM.
Ciudad Universitaria, s/n. 28040 Madrid. Spain.

² CEIGRAM / ETSI Agrónomos. UPM. Ciudad Universitaria, s/n. 28040 Madrid. Spain.

Abstract

Although frost can cause considerable crop damage, and practices have been developed to mitigate forecasted frost, frost forecasting technologies have not changed for years. This paper reports on a new method based on successive application of two models to forecast the number of monthly frost days for several Community of Madrid (Spain) meteorological stations. The first is an autoregressive integrated moving average (ARIMA) stochastic model that forecasts minimum monthly absolute temperature (t_{\min}) and average monthly minimum temperature (μ_t) following Box and Jenkins methodology. The second model relates monthly temperatures (t_{\min} , μ_t) to the minimum daily temperature distribution during one month. Three ARIMA models were identified. They present the same seasonal behaviour (integrated moving average model) and different non-seasonal part: autoregressive model (Model 1), integrated moving average model (Model 2) and autoregressive and moving average model (Model 3). The results indicate that minimum daily temperature (t_{\min}) for the meteorological stations studied followed a normal distribution each month with a very similar standard deviation through out the years. This standard deviation obtained for each station and each month could be used as a risk index for cold months. The application of Model 1 to predict minimum monthly temperatures produced the best frost days forecast. This procedure provides a tool for crop managers and crop insurance companies to assess the risk of frost frequency and intensity, so that they can take steps to mitigate frost damage and estimate the damage that frost would cause.

Additional key words: ARIMA, frost days, mean square error, time series.

Resumen

Predicción de días de helada a partir de temperaturas mensuales

Aunque las heladas pueden causar considerables daños a los cultivos y existen prácticas que mitigan las heladas, las técnicas utilizadas no han cambiado en muchos años. Este artículo proporciona un nuevo método para predecir el número de días de heladas para varias estaciones meteorológicas en la Comunidad de Madrid (España) basado en la aplicación sucesiva de dos modelos. El primero es un modelo estocástico, autorregresivo integrado de media móvil (ARIMA), que predice la temperatura mínima absoluta mensual (t_{\min}) y la temperatura media de mínimas mensual (μ_t) siguiendo la metodología de Box and Jenkins. El segundo modelo relaciona las temperaturas mensuales (t_{\min} , μ_t) con la distribución de temperaturas mínimas diarias de un mes. Se identificaron tres modelos ARIMA. Todos presentan el mismo comportamiento estacional (modelo de media móvil diferenciado) y diferente no estacional: modelo autorregresivo (Modelo 1), modelo de media móvil diferenciado (Modelo 2) y modelo autorregresivo y de media móvil (Modelo 3). Los resultados indican que las temperaturas mínimas diarias (t_{\min}) siguen una distribución normal con una desviación estándar similar a lo largo de los años. Esta desviación estándar podría utilizarse como índice de riesgo para los meses fríos. La aplicación del Modelo 1 para predecir temperaturas mínimas mensuales mostró la mejor predicción en días de helada. Este procedimiento proporciona una metodología para prevenir los daños por heladas en cosechas y estimar el incremento en los daños cuando aparece un escenario inesperado, siendo útil para los agricultores y para las compañías de seguros agrarios.

Palabras clave adicionales: ARIMA, días de helada, error cuadrático mínimo, series de tiempo.

Abbreviations used: ACF (autocorrelation function), ACFP (partial autocorrelation function), AR (autorregresivo), ARIMA (autoregressive integrated moving average), B-J (Box and Jenkins), FD (frost days monthly), MA (moving average), STD (standard deviation).

* Corresponding author: maite.castellanos@upm.es

Received: 24-07-08. Accepted: 03-06-09.

Introduction

Crop establishment and development are crucial because of the investment of time and resources involved. Failure of establishment is mainly due to the influence of climatic factors that cannot be controlled (Penning and Van Laar, 1982; Hutchinson, 1991). Unfortunately, practices to mitigate potential frost damage generally increase crop costs (Hansen *et al.*, 1999). Since freezing temperatures restricting the length of the growing season are responsible for reductions in yield and quality of agricultural crops (Harker, 2002; Faubion, 2003), then minimum temperatures become critical.

Many studies have used monthly temperature, rain, radiation, etc. data to develop models that simulate agro-climatic scenarios (Kuehl *et al.*, 1976; Andersen *et al.*, 2001). Temperature is one of the most important variables in these simulations. Studies with regression functions by Cao and Moss (1989), Jamieson *et al.* (1995) and Landau *et al.* (2000), among others, highlight the relevance of temperature in crop growth. Several temperature models have been developed in order to simulate crop response to daily maximums and minimums of one specific region (Jamieson *et al.*, 1995), or with lower accuracy, response to monthly temperatures (Nonhebel, 1993; Castellanos, 1997) or annual oscillations (Fernández, 1992).

Stochastic simulation of weather data to forecast crop yield has become of great importance (Bannayan and Crout, 1999). Dionne *et al.* (2002) describe crop damage, due to frost, related to crop phenological phase and low temperature tolerance. Knowledge of how temperatures affect crops is basic to selecting and managing crops to avoid risks from extreme temperatures that would seriously reduce yield and market offer (Yang and Chen, 1989; Zhang *et al.*, 1991; Harker, 2002). However, there are very few studies in a given area modeling temperature in relation to frost days as a method to prevent crop damage (Staggenborg *et al.*, 1999).

Air temperature, a temporal series, can be modelled using various techniques, among them the autoregressive integrated moving average (ARIMA) models (McMichael and Hunter, 1972; Kantz and Schreiber, 1997; Montgomery and Zarnowitz, 1998). The aim of this modelling approach is to express current time series values as a linear function of past time series values (the autoregressive component) and current and lagged values of a white noise process (moving average compo-

nent) (Box and Jenkins, 1970). In other words, the aim of the ARIMA models is to separate the observed elements into two components: the first, related to the organized part (including tendency, seasonality, cycles); and the second, the random residuals or white noise (García-Barrón and Pita, 2004). For example, Persaud and Chang (1985) used this type of model to relate maximum and minimum daily temperatures with solar radiation.

ARIMA modelling is applicable to stationary series time. The mean, variance and autocorrelation of a stationary series are independent of time; otherwise a series is non-stationary. A non-stationary series may be transformed into a stationary series by differentiating consecutive values ($W_i = Y_i - Y_{i-1}$) and/or by differentiating consecutive seasonal values if the series is seasonal ($W_i = Y_i - Y_{i-s}$).

ARIMA models could be (p,d,q), or (p,d,q) (P,D,Q)_s, if the series is seasonal, where p and P are the orders of non-seasonal and seasonal autoregressive parameters, q and Q are the orders of non-seasonal and seasonal moving average parameters, and d and D are the numbers of regular and seasonal differences required (as explained above), respectively. Models constructed in this way have been widely used in economics, engineering, etc to explain the time structure of each series and predict its change. The theoretical basis of the proposed models, including required conditions, equations and significance tests, will be explained in the next section.

The aim of this study is to apply these concepts to estimate monthly frost days (FD) based on seasonal ARIMA model forecasted minimum monthly (absolute, t_{\min} and average, μ_t) temperatures. This method allows us to estimate frost risk during months where crops could be damaged due to low temperatures. In Comunidad de Madrid this scenario is observed from February to May and from September to December. In addition, no daily temperature series, only monthly minimum series, easier to obtain from meteorological stations, are required.

Material and methods

Data

Meteorological data, including average monthly minimum temperature (μ_t), minimum monthly absolute temperature (t_{\min}) and number of frost days monthly (FD) corresponding to the study years (1950-2003)

were obtained from twelve meteorological stations in Comunidad de Madrid (Fig. 1) at 3°-4° longitude, 40°-41° latitude and 523-1000 m altitude (Table 1). Data missing from the temperature time series at each meteorological station were filled in by linear regression between the monthly values from neighbouring stations (Guerra-Gómez, 1985), checking the homoscedasticity of the series (Buendía, 1985; Sobrino *et al.*, 2003) previously.

The National Meteorological Institute (INM) and the Regional Meteorological Center of Retiro provided, for each station, the minimum daily temperature in Celsius (t_{dmin}). From this the following were calculated: absolute minimum monthly temperature (t_{min}), average minimum monthly temperature (μ_t) and number of days monthly with $<0^\circ\text{C}$ (FD).

Data for the last 12 months, corresponding to year 2003, were selected to verify the forecasted set of FD, t_{min} and μ_t .

Modelling temperature series

Monthly temperature series has a certain random nature, with an evident seasonal component. Fig. 2 shows the actual data used in the model-building process, and illustrates the high degree of non-linearity and seasonality in time.

The univariant analysis of time series developed by Box and Jenkins (1970) (B-J method) was applied to all

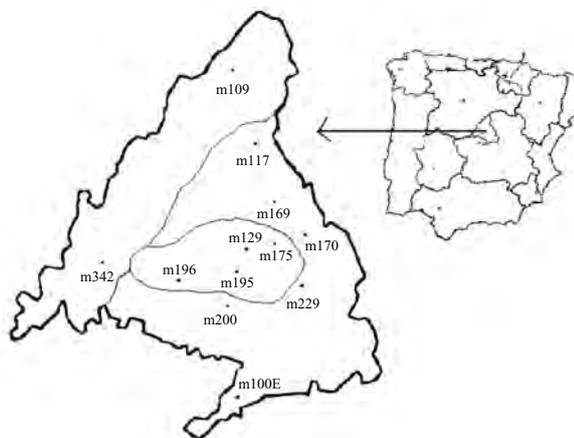


Figure 1. Map of Spain, with enlarged Comunidad of Madrid. The twelve meteorological stations selected are indicated by a filled circle. The urban area in the central part and the mountain area in the north-west, are bounded by a line.

the meteorological stations with data from 1950 till 2003. Before starting with this methodology it is essential to obtain stationary series (Brockwell and Davis, 1987). Normally the differences of the original series remove the non-stationary features of a time series, such as the time-varying mean and variance. For most economic time series, the values of d and D (mentioned in the introduction) are 0 or 1. Preliminary data analysis found that this is also the case for all time series considered in this work. To determinate the values of D and d empirically, tests of Franses (1991) and Beaulieu and Miron (1993) were applied.

The B-J method consists essentially of three steps: a) identification of possible models, b) model parameter estimation and c) comparison of forecast versus data. Each of these steps will be explained below (for more detail see Bowerman and O'Connell, 1987; and Brockwell and Davis, 1987).

Identification of possible models

At this stage we also need to decide how many autoregressive (p) and moving average (q) parameters

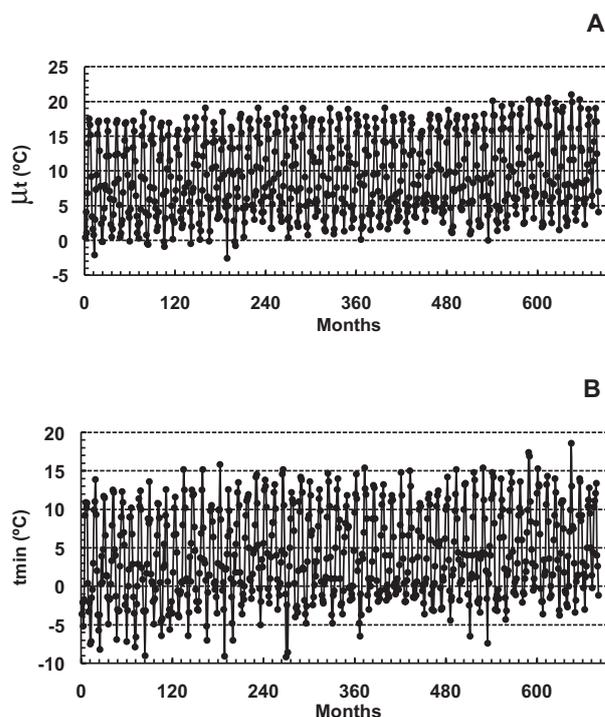


Figure 2. Original series of absolute minimum monthly temperatures (A) and average minimum monthly temperatures (B). Data from station m196, from December 1950 until 2003.

Table 1. Location of the 12 meteorological stations used in this study (longitude, latitude and altitude)

	m100E	m109	m117	m129	m169	m170	m175	m195	m196	m200	m229	m342
Longitude	3°38'	3°38'	3°30'	3°32'	3°18'	3°17'	3°27'	3°40'	3°47'	3°43'	3°18'	4°14'
Latitude	40°01'	40°59'	40°44'	40°27'	40°30'	40°31'	40°29'	40°24'	40°22'	40°18'	40°14'	40°19'
Altitude	490m	1000m	654m	582m	613m	610m	611m	667m	687m	617m	592m	523m

are necessary to yield an effective but still parsimonious model of the process (parsimonious means that it has the fewest parameters and greatest number of degrees of freedom among all models that fit the data). In practice, the numbers of the p or q parameters very rarely need to be greater than 2.

The identification step is based on the behaviour of two functions: the simple autocorrelation function, ACF, and the partial autocorrelation function, ACFP (Aguirre, 1994).

ACF gives the correlation strength between two values of the series separated by *k* values [lag (*k*)], including the intermediate values (Box and Jenkins, 1976). ACF values, between +1 and -1, are calculated as:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{cov}(W_i, W_{i-k})}{\sqrt{\text{var}(W_i) \text{var}(W_{i-k})}} \text{ with } \begin{matrix} k = 0, 1, 2, \dots, N/2 \\ i = 1, 2, \dots, N-k \end{matrix} \quad [1]$$

N is the number of data in the series, *i* is the variable value at a certain month, \bar{W} is the series average, and its estimator is:

$$r_k = \frac{\sum_{i=1}^{N-k} (W_i - \bar{W})(W_{k+i} - \bar{W})}{\sum_{i=1}^N (W_i - \bar{W})^2} \quad [2]$$

ACFP gives the correlation strength between two values of the series separated by *k* values [lag (*k*)], but not the intermediate ones. It is estimated as (Bras and Rodriguez-Iturbe, 1993):

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \quad k = 2, 3, \dots$$

$$r_{kk} = 1 \quad \text{for } k = 1 \quad [3]$$

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \quad j = 1, 2, \dots, k-1$$

Normally a significant interval, 1.92 or 1.96 times the standard error, is calculated for both functions to evaluate if the values obtained are significant (Hamilton, 1994).

The general form of the seasonal ARIMA model (Box and Jenkins, 1970) with periodicity *s* (*s*=12 for monthly data in our case) can be written as:

$$\Delta_s^D \Delta^d W_i = [\phi_1 \Delta_s^D \Delta^d W_{i-1} + \dots + \phi_p \Delta_s^D \Delta^d W_{i-p}] + [\theta_1 \Delta_s^D \Delta^d a_{i-1} + \dots + \theta_q \Delta_s^D \Delta^d a_{i-q} + \Delta_s^D \Delta^d a_i] + [\Phi_1 \Delta_s^D \Delta^d W_{i-s} + \dots + \Phi_P \Delta_s^D \Delta^d W_{i-Ps}] + [\Theta_1 \Delta_s^D \Delta^d a_{i-s} + \dots + \Theta_Q \Delta_s^D \Delta^d a_{i-Qs}] + \delta \quad [4]$$

where δ is white noise and a_i is white noise referred to that month. *B* is the backward shift operator, $\Delta^d = (1-B)^d$ and $\Delta_{12}^D = (1-B_{12})^D$ are, respectively, the operators for the *d*th - order monthly difference and the *D*th - order annual difference. For instance, for *D* = 1, *d* = 1:

$$\Delta W_i = Y_i - Y_i \rightarrow W_i \quad [5]$$

$$\Delta_{12} W_i = Y_i - Y_{i-12} \rightarrow W_i \quad [6]$$

The rest are polynomial terms that involve several unknown parameters to be estimated ($\phi_1 \dots \phi_p, \theta_1 \dots \theta_q, \Phi_1 \dots \Phi_P, \Theta_1 \dots \Theta_Q$). The first two polynomials in Eq. [4] represent, respectively the AR (autorregresive) and MA (moving average) components. The latter two are the seasonal AR and MA polynomials respectively. These components are used to model the cyclical behaviour of a time series around the time-invariant mean.

Estimation of model parameters

Estimation of model parameters obtained in the previous section ($\phi_1 \dots \phi_p, \theta_1 \dots \theta_q, \Phi_1 \dots \Phi_P, \Theta_1 \dots \Theta_Q$) is based on the squared unconditioned minimums (Harvey, 1990). Later, model selection is based on parameters and residual tests such as standard error and t-Student test. Finally, the residual series, the difference between series estimates and observable values was analyzed to check that it was indeed a random series (Harvey and Pierse, 1984), that is, with a null average and normally distributed.

The model selected will be the one that minimizes the mean - square error (MSEt) in the series estimation (Uriel, 1995) defined as:

$$MSE_t = \sum_{j=1950}^{2002} \sum_{i=1}^{12} \frac{(\hat{t}_{ij} - t_{ij})^2}{N} \quad [7]$$

where: \hat{t}_{ij} is the estimated monthly temperature, t_{ij} is the observed monthly temperature, i = month number, j = year number, N = number of data.

Comparison of forecast with data

Once the simulation is finished, a model has to be selected from the possible ones based on accuracy according to B-J methodology.

The forecast for the last year (2003) with the selected model for each series based on data from 1950 to 2002 were made, with a confidence interval of 95%. The error estimated in this forecast is calculated based on Eq. [7] restricted to year 2003.

When forecast values are within a certain confidence limit of the real data, the models can be considered adequate to forecast the series for the following year (Bowerman and O'Connell, 1987).

Forecast of temperature series

The estimation process is performed on transformed (differenced) data; before the forecasts are generated, the series needs to be integrated (integration is the inverse of differencing) so that the forecasts are expressed in values compatible with the input data. This automatic integration feature is represented by the letter I in the name of the methodology ARIMA.

At this step, the time series (from 1950 till 2003) is used to adjust model parameters. Once finished, the model is used to forecast the following year.

Estimation of number of frost days

A normal probability plot was constructed for each month and each station to check that the underlying probability distribution of minimum daily temperature can be assumed to be normal (Eq. [8]).

$$f(t_{\min}) = \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{(t_{\min} - \mu_t)^2}{2\sigma_t^2}} \quad [8]$$

The essence of such a plot is that points plotted from a normal distribution fall close to a straight line

(Devore, 1995). From these linear regressions the mean and standard deviation can be calculated.

Once that normality was checked, we can assume that the absolute minimum monthly temperature (t_{\min}) will be equal to the average minimum monthly temperature (μ_t) minus a coefficient (α) related through the typical deviation (σ_t). That is (Tarquis *et al.*, 1993):

$$t_{\min} = \mu_t - \alpha\sigma_t \quad \text{or} \quad \alpha = \frac{\mu_t - t_{\min}}{\sigma_t} \quad [9]$$

Applying Eq. [9], the coefficient was estimated for each station and each month.

Estimated FD, the probability of a minimum daily temperature lower than 0°C (Diehl *et al.*, 1982), depends on the parameters of the normal distribution, μ_t and σ_t .

$$Pr(t_{\min} < 0) = \frac{1}{\sigma_t \sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{t^2}{2\sigma_t^2}} dt \quad [10]$$

$$\text{where } \sigma_t = \frac{\mu_t - t_{\min}}{\alpha}$$

Results and discussion

Modelling temperature series

Once that all series have been transformed into stationary series, ACF and ACFP graph were studied (see Fig. 3). There were three possible models for the non-seasonal part and only one model for the seasonal part, depending on the station and temperature analyzed. These are specific cases of the general model shown in Eq. [4].

- Model 1: autoregressive model of order one in the non-seasonal part, and integrated moving average model of order one for the seasonal part. In B-J methodology this would be: (1,0,0),(0,1,1)₁₂

$$\Delta_{12}W_i = \phi\Delta_{12}W_{i-1} + \Delta_{12}a_i + \theta \Delta_{12}a_{i-1} + \Theta\Delta_{12}a_{i-12} \quad [11]$$

- Model 2: integrated moving average model of order one in the non-seasonal part, and integrated moving average model of order one for the seasonal part, (0,1,1),(0,1,1)₁₂:

$$\Delta_{12}\Delta W_i = \theta\Delta_{12}\Delta a_{i-1} + \Delta_{12}\Delta a_i + \Theta\Delta_{12}\Delta a_{i-12} \quad [12]$$

- Model 3: autoregressive and moving average model of order one in the non-seasonal part, and integrated moving average model of order one for the seasonal part, (1,0,1),(0,1,1)₁₂ :

$$\Delta_{12}W_i = \phi\Delta_{12}W_{i-1} + \Delta_{12}a_i + \theta \Delta_{12}a_{i-1} + \Theta\Delta_{12}a_{i-12} \quad [13]$$

For both series, and , Model 1 is the most frequently fitted (Table 2). Model 2 μ_t time series only fitted two stations (m109 and m342), a water reservoir located in mountain area of Madrid and a hydroelectric power station in a valley, respectively. This model has been fitted for more unstable series than those fitted with Model 1, in which it was not necessary to differentiate the non-seasonal part. These two stations frequently experience temperature inversions due to cold air accumulation in the valleys. The t_{min} series fitted two stations (m109 and m195), a water reservoir (m109) and the city centre surrounded by forest and vegetation (m195), a differentiating it from other stations in the same area. Finally, both time series fitted the station located at Barajas airport, with its higher instability due to winds and turbulence created by aeroplanes.

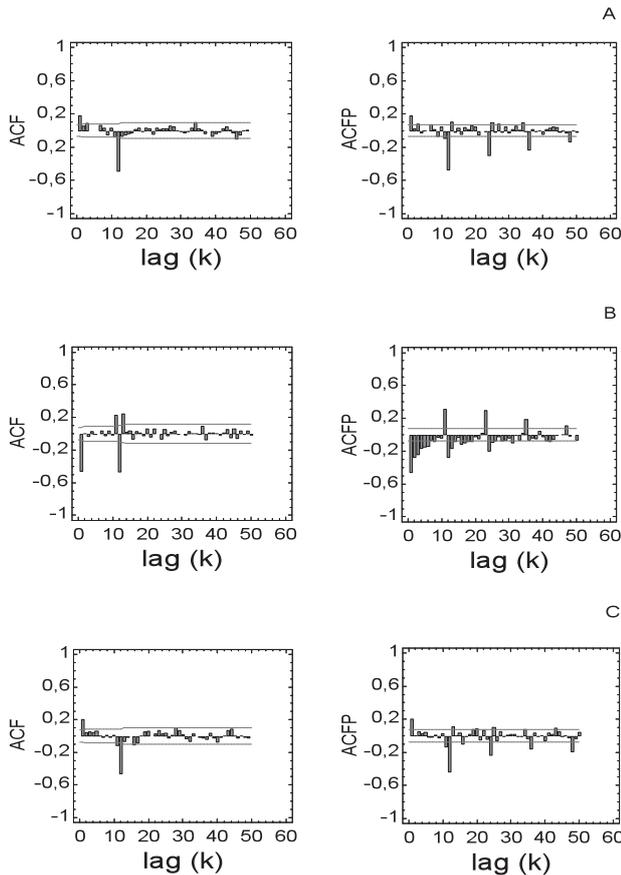


Figure 3. Simple Autocorrelation Function (ACF) and Partial Autocorrelation function (ACFP) of various ARIMA models identified : A) ARIMA(1,0,0)(0,1,1)₁₂ from station m196, B) ARIMA(0,1,1)(0,1,1)₁₂ from station m109 and C) ARIMA(1,0,1)(0,1,1)₁₂ from station m129. The series used were differentiated average minimum monthly temperatures.

Table 2. ARIMA models fitted for minimum temperature series per station: monthly absolute (t_{min}) minimum and monthly average minimum (μ_t)

Stations code	ARIMA models (s=12)	
	μ_t	t_{min}
m100E	(1,0,0), (0,1,1) _s	(1,0,0), (0,1,1) _s
m109	(0,1,1), (0,1,1) _s	(0,1,1), (0,1,1) _s
m117	(1,0,0), (0,1,1) _s	(1,0,0), (0,1,1) _s
m129	(1,0,0), (0,1,1) _s	(1,0,1), (0,1,1) _s
m169	(1,0,0), (0,1,1) _s	(1,0,0), (0,1,1) _s
m170	(1,0,0), (0,1,1) _s	(1,0,0), (0,1,1) _s
m175	(1,0,0), (0,1,1) _s	(1,0,0), (0,1,1) _s
m195	(1,0,0), (0,1,1) _s	(0,1,1), (0,1,1) _s
m196	(1,0,0), (0,1,1) _s	(1,0,0), (0,1,1) _s
m200	(1,0,0), (0,1,1) _s	(1,0,0), (0,1,1) _s
m229	(1,0,0), (0,1,1) _s	(1,0,0), (0,1,1) _s
m342	(0,1,1), (0,1,1) _s	(1,0,0), (0,1,1) _s

The ARIMA nomenclature correspond to (p,d,q) (P,D,Q)s, where p and P are orders of the non-seasonal and seasonal autoregressive parameters, q and Q are orders of the non-seasonal and seasonal moving-average parameters, d and D are numbers of regular and seasonal differences required, and s is the periodicity.

Coefficient values of autoregressive Model 1 with an average of 0.4 indicate a low correlation between variable value (Tables 3 and 4). Moving average Model 2 indicates that the possible tendency that could exist is due to white noise. Coefficient values of an average of 0.86 (Tables 3 and 4) indicate considerable influence of preceding and present noise.

The seasonal part is always an integrated moving average model that yields a random and very weakly auto-correlated series type. Coefficients of this part, of an average of 0.90 (Tables 3 and 4), indicate considerable influence of white noise between consecutive years.

Forecast of temperature series

Forecast temperatures for the twelve stations were well adjusted. An average forecast was calculated for absolute temperatures and minimum averages for the last year according to the chosen confidence interval (95%). Absolute minimum temperature average error was 1.3°C, with a standard deviation (STD) error of 1.03. Minimum average temperature errors averaged of 1.1°C, with STD of 0.64. It is observed in general that, for the

Table 3. Coefficient values of models obtained per station for monthly average minimum temperature (μ_t)

Meteorological Stations	Coefficients ^a		
	ϕ	θ	Θ
m100E	0.424		0.926
m109		0.921	0.956
m117	0.502		0.938
m129	0.676	0.451	0.964
m169	0.908		0.950
m170	0.327		0.964
m175	0.187		0.961
m195	0.238		0.917
m196	0.290		0.906
m200	0.293		0.948
m229	0.411		0.843
m342		0.877	0.959

^a Coefficients correspond to the following models:

$$(1,0,0)(0,1,1)_{12} \quad \Delta_{12}W_t = \phi \Delta_{12}W_{t-1} + \Delta_{12}a_t + \Theta \Delta_{12}a_{t-12}$$

$$(0,1,1)(0,1,1)_{12} \quad \Delta_{12}\Delta W_t = \theta \Delta_{12}\Delta a_{t-1} + \Delta_{12}\Delta a_t + \Theta \Delta_{12}\Delta a_{t-12}$$

$$(1,0,1)(0,1,1)_{12} \quad \Delta_{12}\Delta W_t = \theta \Delta_{12}\Delta a_{t-1} + \Delta_{12}\Delta a_t + \Theta \Delta_{12}\Delta a_{t-12}$$

twelve stations, the average minimum temperatures adjust better than the absolute minimums (Figs. 4 and 5).

As an example, some data with actual temperatures for three stations are shown in Figs. 4 and 5, each one with a different model. In these Figures the good adjustment of estimated to actual temperatures can be observed. In all months, average forecasted temperatures are not outside the limits of the 95% confidence interval. For m109, as well as for the rest of the stations modelled by Model 2, there is not a significant difference between the model forecast and the monthly average of the series, confirming that this model only reflects the influence of white noise.

Estimation of number of frost days

Estimation of number of frost days monthly (FD) for two crop periods with highest risk per month and per station from the years studied is shown in Table 5.

December is the month with the highest FD at all stations, followed by February and November. From February to May, stations m109 and m169 show the highest accumulated FD value, while from September to December station m100E does so, in agreement with their geographic situation (Fig. 1).

Table 4. Coefficient values of the models obtained per station for monthly absolute minimum temperature (t_{min})

Meteorological Stations	Coefficients ^a		
	ϕ	θ	Θ
m100E	0.133		0.913
m109		0.938	0.939
m117	0.602		0.921
m129	0.751	0.576	0.955
m169	0.556		0.942
m170	0.308		0.882
m175	0.200		0.956
m195		0.953	0.937
m196	0.088		0.937
m200	0.140		0.966
m229	0.456		0.919
m342	0.523		0.934

^a Coefficients: see Table 3.

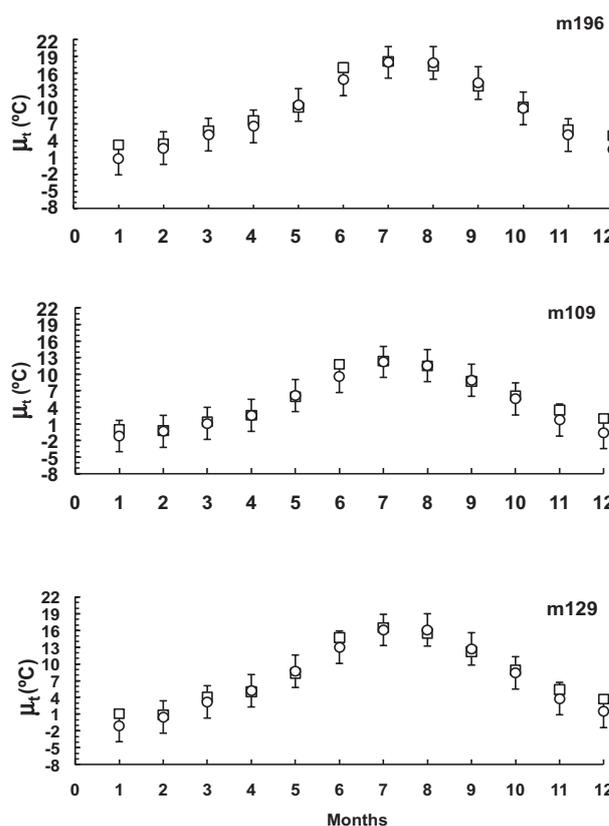


Figure 4. Average minimum monthly temperature series at 2003: original data represented by an empty square (\square) and forecasted represented by an empty circle (\circ), for stations m196, m109, m129. Error bars correspond to the selected confidence interval (95%) in the forecast.

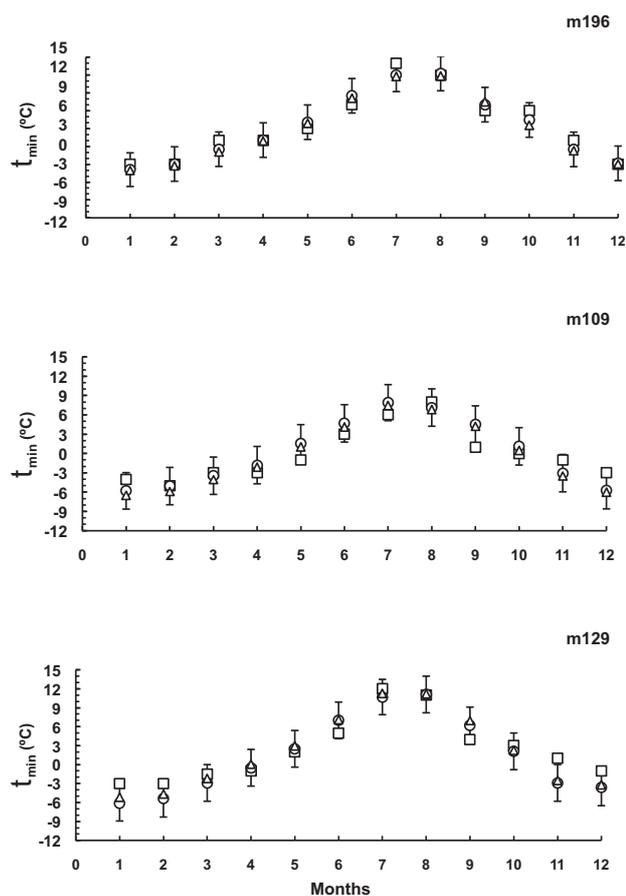


Figure 5. Absolute minimum monthly temperature series in 2003: original data represented by an empty square (\square) and forecasted, represented by an empty circle (\circ) for stations m196, m109, m129. Error bars correspond to the selected confidence interval (95%) in the forecast.

The same study can be carried out on the STD of FD (see Table 6). One of the first results is that FD standard values follow the same pattern explained earlier for each FD value. Therefore, the month with the highest variation of FD is December, followed by February and then March and November with a similar STD.

The r^2 values obtained from the probability plot analysis (Fig. 6) are greater than 89% for all cases, confirming that a normal distribution can be used to describe daily minimum temperatures. The range of estimated values varies from 4.2 to 0.204 (Table 7). In general, comparing Tables 5 and 6, an inverse relation between α and the average number of monthly frost days can be seen. The lack of any clear relation between forecasted coefficient errors and location of stations is an indication of the complexity due to several simultane-

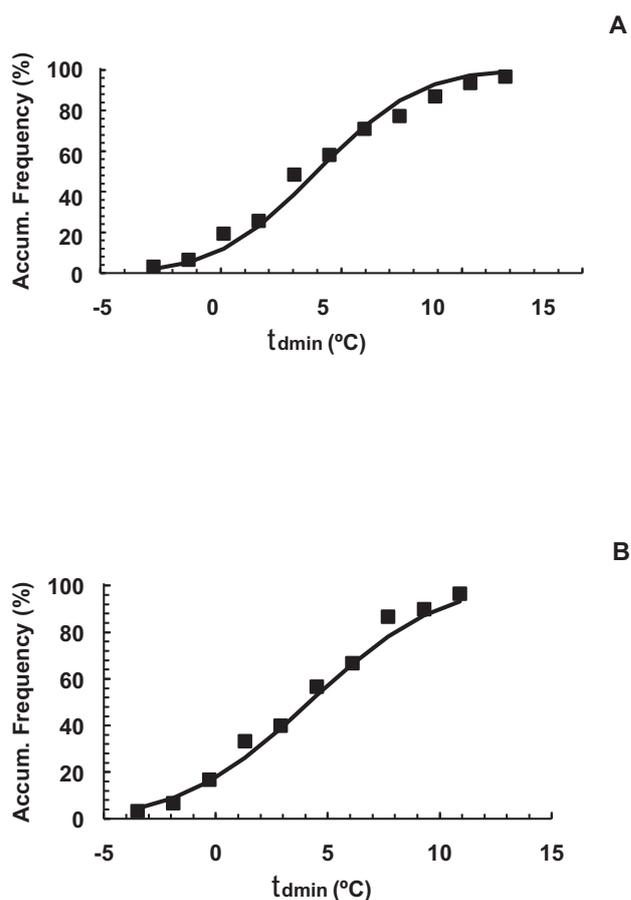


Figure 6. Accumulated normal probability plot based on daily minimum temperatures (t_{dmin}) for two different months at station m100E. Actual data represented by a filled square (\blacksquare) and the fitted normal probability is represented by a continuous line: A) April 1995 average 3.99, STD 3.3 and $r^2=0.98$ B) December 1995 average 4.18, STD 4.5 and $r^2=0.99$.

ous influencing factors (water reservoirs, urban centre, mountain area, etc).

The number of monthly frost days is calculated using forecasted temperatures (t_{\min} and μ_t) and estimated σ_t and α , as explained previously.

Some of the forecasted frost days are shown in Fig. 7, where for each month studied the monthly average and standard deviation of FD are also presented. For the three stations of Fig. 7, the highest error is in February and November, as in the rest of the stations analyzed.

To evaluate the adjustment of forecasted to actual data, a coefficient error is calculated to compare the difference between both and the monthly standard deviation found (Table 8). We can observe that almost all errors are smaller than 0.5 STD. Only twelve values are larger than STD, mainly in February and November.

Table 5. Estimation of number of monthly frost days per station at various periods: monthly, from February to May (total I), from September to December (total II) and sum of the months shown (total I+II)

Month	Meteorological Stations											
	m100E	m109	m117	m129	m169	m170	m175	m195	m196	m200	m229	m342
February	14	17	15	13	17	14	9	5	7	9	16	3
March	12	13	10	7	12	10	5	2	3	3	9	1
April	4	5	4	1	5	4	1	0	1	1	4	0
May	0	1	0	0	1	1	0	0	0	0	1	0
September	0	0	0	0	0	0	0	0	0	0	0	0
October	2	1	1	0	2	1	0	0	0	0	1	0
November	12	11	8	7	11	9	4	2	3	4	10	1
December	19	17	17	15	18	15	12	6	8	13	16	4
Total I	31	37	28	21	34	29	16	8	11	13	29	4
Total II	32	29	25	22	30	25	16	7	11	17	27	5
Total I+II	63	66	53	43	64	54	32	15	22	30	56	9

Table 6. Standard deviation (STD) values of the number of monthly frost days (FD)

Month	m100E	m109	m117	m129	m169	m170	m175	m195	m196	m200	m229	m342
February	7.59	4.54	6.01	5.52	5.63	5.54	5.52	5.55	5.03	4.72	6.19	3.66
March	5.21	6.24	4.53	4.62	6.64	5.94	4.00	2.69	3.53	3.15	6.24	2.36
April	3.41	3.54	2.79	1.65	3.38	3.23	1.94	0.77	1.21	0.92	2.72	0.00
May	0.34	1.24	0.67	0.24	1.54	0.88	0.36	0.00	0.36	0.48	0.88	0.00
September	0.00	0.17	0.00	0.18	0.34	0.26	0.17	0.00	0.00	0.00	0.41	0.00
October	2.67	2.46	1.30	0.81	2.60	1.61	0.65	0.13	0.24	0.74	2.28	0.00
November	5.70	6.37	5.17	4.51	6.13	5.60	3.91	2.14	3.38	3.29	6.06	1.54
December	7.60	5.80	7.12	7.29	8.21	7.61	6.89	5.14	6.19	6.95	9.13	4.28

Table 7. α coefficient values estimated per month and per meteorological station

Month	m100E	m109	m117	m129	m169	m170	m175	m195	m196	m200	m229	m342
February	1.300	0.600	0.400	0.204	1.000	0.400	1.000	1.63	1.52	1.501	0.500	1.344
March	0.745	0.900	1.579	1.149	1.013	1.184	1.300	1.60	1.67	1.794	0.600	1.388
April	1.413	1.571	1.596	1.770	1.518	1.667	1.741	1.72	1.78	1.875	1.690	4.200
May	1.739	1.668	1.741	2.032	1.608	1.799	1.785	4.20	1.80	1.644	1.623	2.600
September	2.600	1.682	4.200	1.577	1.472	1.665	1.484	2.60	2.60	2.600	1.564	1.800
October	1.440	1.556	1.664	1.832	1.476	1.434	1.902	1.78	1.89	1.612	1.500	2.600
November	0.800	1.074	1.362	1.265	1.300	0.800	1.435	1.58	1.46	1.466	1.213	1.453
December	0.900	1.200	1.400	1.200	0.560	0.342	1.504	1.42	1.40	0.629	0.600	1.277

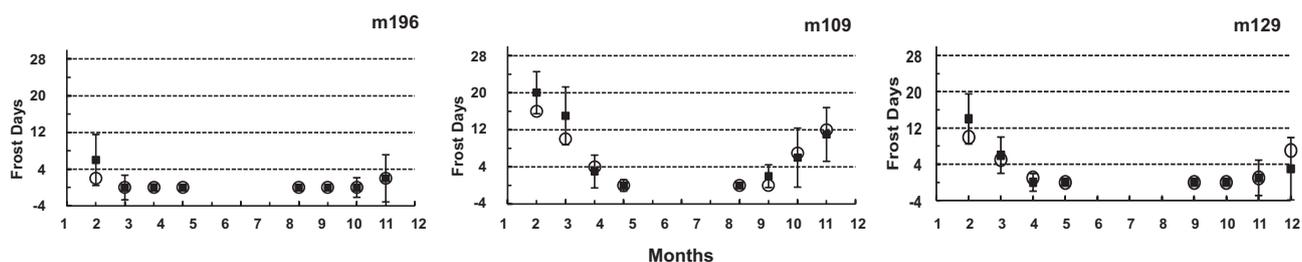


Figure 7. Series of number of monthly frost days series, 2003: original data represented by a filled square (■) and forecasted, represented by an empty circle (O) for stations m196, m109, m129, for two crop periods with highest frost risk in Comunidad de Madrid, from February to May and from September to December, with error bars corresponding to the monthly standard deviation.

Observing all the stations, there are only three that present an error larger than 1.5 STD in one month (m109 and m342), and in two stations all months show less than 0.6 STD (m196 and m195).

Comparing the results given in Figs. 6 and 7, it is obvious that the increase of the forecasted error occurs when the average minimum monthly temperature is very close to 0°C. In this case, an error in the forecasted temperature of 1°C varies the number of frost days much more than the same error in another month where the value is farther from the 0°C threshold value. On the other hand, when the temperature series follows Model 2, almost the same forecast can be obtained with average monthly temperature.

Conclusions

Application of B-J methodology to analyze and predict temporary monthly temperature series is effective. The twelve Community of Madrid stations studied yielded three different models: $(1,0,0)(0,1,1)_{12}$ (Model 1),

$(0,1,1)(0,1,1)_{12}$ (Model 2) and $(1,0,1)(0,1,1)_{12}$ (Model 3). Elevated values of the moving-average coefficient in the seasonal part of the three models indicate influence of noise in the seasonal series (from year to year). In general, the forecast for minimum temperatures is in good agreement with the data. However, Model 1 series yielded better results than Model 2 series.

The estimation of FD starting from the minimum absolute temperatures and the monthly average of minimums, assuming a normal distribution of minimum daily temperatures, is an effective method. Of the twelve stations studied, the forecast for eleven are very close to the actual values. The estimated FD for February and October show the higher forecasted errors, while March, April and May were much better adjusted. The reason for this is the size of the forecasted error of the average minimum temperature when close to 0°C. In that situation the error generated in the forecasted frost days increases, more than the error in the value of the a parameter.

This study indicates that it is possible to transform temperatures into a risk index for various crops, due to the possibility of calculating the number of days below

Table 8. Monthly coefficient error between forecast and monthly real data corresponding to year 2003

Month	m100E	m109	m117	m129	m169	m170	m175	m195	m196	m200	m229	m342
February	0.66	1.76	0.50	0.54	1.42	1.26	1.09	0.54	0.20	0.85	1.45	1.64
March	0.19	0.80	0.22	0.43	0.15	0.17	0.25	0.00	0.28	0.32	0.48	0.85
April	0.29	0.00	1.08	0.00	0.89	0.62	0.52	0.00	0.00	1.09	0.37	0.00
May	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
September	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
October	0.00	0.81	0.00	1.23	0.38	1.24	0.00	0.00	0.00	0.00	0.88	0.00
November	0.53	0.00	0.19	0.22	0.65	0.18	0.00	0.00	0.30	0.30	0.17	0.65
December	1.32	0.00	0.56	0.69	0.85	0.39	0.58	0.00	0.48	1.15	0.22	0.47

a certain temperature, not only 0°C. Although this methodology was carried out in the Community of Madrid area, it may be applied to other locations where the minimum daily temperature (t_{dmin}) can be assumed to be normally distributed.

Acknowledgments

This research was supported by Comunidad de Madrid Research Project 076/92. The cooperation of the Spanish National Meteorological Institute and the Spanish Ministerio de Agricultura, Pesca y Alimentación (MAPA) is gratefully acknowledged.

References

- AGUIRRE J., 1994. Introducción al tratamiento de series temporales. Aplicación a las ciencias de la salud. Ed. Díaz de Santos, Madrid. [In Spanish].
- ANDERSEN J.A., ALAGARSWMY G., ROTZ C.A., RIITCHIE J.T., LEBARON A.W., 2001. Weather impacts on maize, soybean, and alfalfa production in the Great Lakes Region, 1895-1996. *Agron J* 93, 1059-1070.
- BANNAYAN M., CROUT N.M.J., 1999. A stochastic modeling approach for real-time forecasting of winter wheat yield. *Field Crop Res* 62, 85-95. doi:10.1016/S0378-4290(99)00008-8.
- BEAULIEU J.J., MIRON J.A., 1993. Seasonal unit roots in aggregate US data. *J Econometrics* 55, 305-328.
- BOWERMAN B.L., O'CONNELL R.T., 1987. Time series forecasting. United concepts and computer implementation. 2nd ed. Duxbury Press, Boxton.
- BOX G.E.P., JENKINS G.M., 1970. Time series analysis: Forecasting and control. San Francisco: Holden Day. doi: 10.1086/295406.
- BOX G.E.P., JENKINS G.M., 1976. Time series analysis: Forecasting and control. 2nd ed. San Francisco: Holden Day. doi: 10.1177-058310248201400608.
- BRAS R.L., RODRIGUEZ-ITURBE I., 1993. Random functions and hydrology. 2nd ed. Addison-Wesley, NY.
- BROCKWELL P.J., DAVIS D.R.A., 1987. Time series: theory and methods. Springer-Verlag, Inc. NY.
- BUENDIA G., 1985. Régimen normal de precipitaciones en la provincia de Valladolid. Instituto Nacional de Meteorología A-102, Madrid. [In Spanish].
- CAO W., MOSS D.N., 1989. Temperature effect on leaf and phyllochron in wheat and barley. *Crop Sci* 29(4),1018-1021.
- CASTELLANOS M.T., 1997. Estudio y predicción de temperaturas mediante análisis de series temporales en la Comunidad de Madrid. PhD Thesis. UPM, Madrid. [In Spanish].
- DEVORE J.L., 1995. Probability and statistics for engineering and the sciences. 4th ed. Duxbury Press, Washington.
- DIEHL R., MATEO BOX J.M., URBANO P., 1982. Fitotecnica general. 2th Ed. Mundi Prensa Ed., Madrid. [In Spanish].
- DIONNE J., CASTONGUAY Y., NADEAU P., DESJARDINS Y., 2002. Freezing tolerance and carbohydrate changes during cold acclimation of green- type annual bluegrass ecotypes. *Crop Sci* 41, 443-451.
- FAUBION D., 2003. Reducing freeze damage with nitrogen. *American Fruit Grower* 123, 40-42.
- FERNANDEZ C.J., 1992. Simulation of normal annual and diurnal temperature oscillation in non-mountainous Mainland United States. *Agron J* 84(2), 244-251.
- FRANSES P.H., 1991. Seasonality, non-stationarity and the forecasting of monthly time series. *Int J Forecast* 7, 199-208. doi: 10.1016/0169-2070(91)90054-Y.
- GARCÍA-BARRÓN L., FERNANDA-PITA M.A., 2004. Stochastic analysis of time series of temperatures in the southwest of the Iberian Peninsula. *Atmósfera* 23, 225-244.
- GUERRA-GÓMEZ J., 1985. Estudio de un modelo estocástico de precipitaciones en la Comunidad Autónoma de Madrid. Aplicaciones. Inst. Nacional Meteorológico, Madrid. A-109. 81 pp. [In Spanish].
- HAMILTON J.D., 1994. Time series analysis. Princeton University Press. Princeton, NJ.
- HANSEN J.W., HONES J.W., KIKER C.F., HODGES A.W., 1999. El Niño-Southern Oscillation impacts on winter vegetable production in Florida. *J Climate* 12, 92-103.
- HARKER K., 2002. Getting the most from a cold forecast. *Florida Grower* 95, 26-28.
- HARVEY A., 1990. The econometric analysis of time series. 2nd ed. Philip Allan, Great Britain.
- HARVEY A.C., PIERCE R.G., 1984. Estimating missing observations in economic time series. *J Am Stat Assoc* 79, 125-131. doi:10.2307/2288346.
- HUTCHINSON M.F., 1991. Climatic analysis in data sparse regions. In: Climatic risk in crop production: models and measurement for the semiarid tropics and arid subtropics (Muchow R.C., Bellamy J.A., eds.). CAB International. Wallingford, pp. 55-73.
- JAMIESON P.D., BROOKING I.R., PORTER J.R., WILSON D.R., 1995. Prediction of leaf appearance in wheat: a question of temperature. *Field Crop Res* 41, 35-44. doi:10.1016/0378-4290(94)00102-I.
- KANTZ H., SCHREIBER T., 1997. Nonlinear time series analysis. Cambridge Univ. Press, Cambridge.

- KUEHL R.O., BUXTON D.R., BRIGGS R.E., 1976. Application of time series analysis to investigate crop and environment relationships. *Agron J* 68, 491-495.
- LANDAU S., MITCHELL R.A.C., BARNETT V., COLLS J.J., CRAIGON J., PAYNE R.W., 2000. A parsimonious, multiple-regression model of wheat yield response to environment. *Agr Forest Meteorol* 101, 151-166. doi:10.1016/S0168-1923(99)00166-5.
- McMICHAEL F., HUNTER J., 1972. Stochastic modeling of temperature and flow in rivers. *Water Resour Res* 8(1), 87-98. doi: 10.1029/WR008i001p00087.
- MONTGOMERY A.L., ZARNOWITZ V., 1998. Time series analysis: Unemployment-United States—forecasting. *Journal Am Stat Assoc* 93, 478-492. doi:10.2307/2670094.
- NONHEBEL S., 1993. The importance of weather data in crop growth simulation models and assessment of climatic change effects. Wageningen University, The Netherlands.
- PENNING F.W.T., VAN LAAR H.H., 1982. Simulation of plant growth and crop production. Pudoc Ed., Wageningen.
- PERSAUD N., CHANG A.C., 1985. Times series analysis of daily solar radiation and air temperature measurements for use in computing potential evapotranspiration. *T ASAE* 28(2), 462-470.
- SOBRINO E., TARQUIS A.M., DÍAZ M.C., 2003. Modeling the oleic acid content in sunflower oil. *Agron J* 95, 329-334.
- STAGGENBORG S.A., VANDERLIP R.L., ROGGEKAMP G.J., KOFOID K.D., 1999. Methods of simulating freezing damage during sorghum grain fill. *J Agron* 91(1), 46-53.
- TARQUIS A.M., SAA A., CASTELLANOS M.T., 1993. Index of freezing based on time series. In: *Plant production on the threshold of a new century* (Struik *et al.*, eds.), Kluwer Acad. Publishers, The Netherlands. pp. 303-310.
- URIEL E., 1995. Análisis de series temporales. Modelos ARIMA. 3ª ed., Colección Abaco, Ed. Paraninfo, Madrid, Spain. [In Spanish].
- YANG X.W., CHEN J.L., 1989. A mathematical model for prediction of time course of vegetable supply based on meteorological factors. *Agr Forest Meteorol* 49, 35-44. doi: 10.1016/0168-1923(89)90060-9.
- ZHANG Y., YU X., GUAN D., 1991. Some applications of syntactic pattern recognition in assessments of temperature conditions in various phenological phases of maize. *Agr Forest Meteorol* 55, 85-93. doi: 10.1016/0168-1923(91)90023-J.