Stochastic analysis of time series of temperatures in the south-west of the Iberian Peninsula

L. GARCÍA-BARRÓN

Departamento de Física Aplicada II, Universidad de Sevilla, España
e-mail: leoncio@us.es

MA. FERNANDA PITA

Departamento de Geografía Física y A.G. R.., Universidad de Sevilla, España

RESUMEN

Se han elaborado los modelos autorregresivos integrados de media móvil (ARIMA) para las series mensuales de larga duración, de temperaturas máximas y mínimas, de los observatorios del suroeste español. Se han transformado las series originales para hacerlas estacionarias. Se han obtenido los órdenes (p, d, q) para cada serie mensual y, tras los procesos de validación por medio de las series de residuos, también se han obtenido los parámetros de las funciones estimadas de los modelos estocásticos mensuales. Como consecuencia, se han formulado predicciones sobre su evolución: las previsiones de los modelos ARIMA inducen a esperar que la temperatura mínima en la región objeto de estudio se encuentra en una fase de incremento térmico medio previsto del orden de 0.2 ºC en el próximo decenio. Sin embargo, no está definida significativamente la evolución de las temperaturas máximas. Estas estimaciones se corresponden con las obtenidas previamente por los autores, en estudios anteriores, por otros métodos, pero con mucha mayor precisión en este caso.

ABSTRACT

Autoregressive Integrated Moving Average (ARIMA) models have been devised for long-term monthly series of maximum and minimum temperatures from south-western Spanish observatories. The original series were transformed into stationary ones, and the orders (p, d, q) for each monthly series were obtained. These were validated using series of residuals, and the parameters of the functions were estimated from the monthly stochastic models. The forecasts of change formulated from the ARIMA models suggest that the minimum temperatures in the study area are in a mean warming phase of the order of 0.2 ºC over the next decade. The change in maximum temperatures is not significantly defined. These estimations correspond with those obtained by the authors in earlier studies using other methods, but the level of precision is much higher in this case.

Key words: Stochastic processes; time series; climate change; Iberian Peninsula.

1. Introduction

The aim of stochastic analysis of time series by the formulation of ARIMA models is to separate the observed elements into two components: The first contains the organized part, including the trend, and the second the random residuals. The models constructed following such separation are used in an attempt to explain the time structure of each series and predict its change. The time series is considered a set of observables of the variables –maximum and minimum temperatures– measured at regular intervals. In this study, temperature measurements made during the period of observation are grouped to form monthly series.

Each of the series must be tested as fulfilling the statistical conditions required by the Box-Jenkins method for such models. We then proceed to identify and estimate the corresponding parameters, so that the estimated model not only fits the observations well, but also enables forecasts to be made.

The theoretical basis of the proposed models, including the definition of conditions of applicability, deduction of equations, and significance of the results, has been set out by various authors (Chatfield, 1992; Box and Jenkins, 1970; Kendall and Ord, 1990; Wei, 1990; Uriel, 1995; Peña, 1995; Ferrand, 1997). Others have used autoregressive moving average techniques in sociology and economics, and also, to a lesser degree, specifically in the study of meteorological variables (Garrido and García, 1994; Al-awadhi and Jolliffe, 1998; Martín *et al.*, 1998; Zekai, 1998, Bravo, *et al.*, 2001; Greiser *et al.*, 2002; Solange and Pinto, 1996).

The following section indicates the observatories used in this study and the characteristics of stationarity required for the time series. Section 3 briefly describes the methodological basis of Autoregressive Integrated Moving Average (ARIMA) processes. Section 4 determines the order for each series, the corresponding functions generated by the model, and the respective residual series. Some of them have been represented graphically as an example. Section 5 estimates, from the results obtained for each monthly series, the forecast for the following decade.

2. Data

The observatories selected for the study are those from south-west Spain (western Andalucía and the south of Extremadura) having data series of the order of a century (Fig. 1). The records are from Sevilla-Tablada (37º 22´N, 6º 00´ W), Córdoba (37º 53´N, 4º 47´W) Badajoz (38º 53´ N, 6º 58´ W), and Huelva (37º 15´, 6º 57´W). The series of variables analyzed are the monthly maximum and monthly minimum temperatures from 1882 to 1999. (As the Huelva observatory ceased activity in 1983, no prediction has been made in section 5 due to the lack of data for the final period).

The data were previously subjected to a process of ''gap-filling'' and tested for uniformity (García-Barrón and Pita, 2001). We consider that these observatories characterize the south-western Iberian Peninsula climate, a region of interest owing to the Azores anticyclone direct action on it. Geo-graphically, the region is delimited by Lisbon and Gibraltar, continental points used by some authors to demonstrate the effects of the North Atlantic Oscillation (Hurrell, 1995; Pita López *et al.* 1999; Pozo-Vázquez *et al.*, 2000).

Fig. 1. Location of the meteorological observatories used in this study.

It is a prerequisite that the time series be stationary before application of the Box-Jenkins method, so this must be verified and, if necessary, appropriate transformations made.

Accepting that the (X_t) series fluctuate around the mean value m that characterizes them, we can express the perturbations δ_t, as the differences between the observed values (x_t) and the μ value,

$$c_t = m + d_t.$$

The process is considered *stationary* if, for any of the subseries (X_{jt}) into which the complete time series is fragmented:

- the mean is constant −that is, the mathematical expectancy of the perturbation is zero.
- the variance is constant; in such case, the perturbation is denominated homocedastic.
- the perturbations are independent of each other −that is, the correlation r_k between two observations, x_t and x_{t+k}, depends only on the number of lags separating them.

Furthermore, it is a requirement that the perturbations have a normal distribution, explicable by the central limit theorem, such that they cannot be imputed to specific causes but result from the effect of multiple random factors.

In the previous studies carried out (García-Barrón, 2000), we obtained the following results:

- Regarding the stability of the mean, most of the monthly series show a trend in the minimum temperatures (significant and with a slope coefficient between 1.7º and 3.3º per century, depending on month and observatory) and also in the maximum temperatures (although not constant in the sign of the slope, and with no statistical significance in some cases). It is therefore recommendable that all the series undergo first-order differencing to eliminate this trend (owing to the stability in variance a second derivation is not suggested).

• Regarding variance, the Levenne test shows that a great number of original series remain within the limits of stability, with a quasi-normal distribution. Nevertheless, logarithmic transformation generally improves variance, in particular in those series with homocedasticity not initially accepted.

Thus, to unify the procedure, the adaptation criterion for all the series will be a first-order linear derivation and logarithmic transformation. The Kolmogorov-Smirnov test confirms that the transformed monthly series fit the Gaussian distribution and can be accepted as meeting the requirement of normality.

Save the exceptions indicated below, the monthly series of maximum and minimum temperatures from the selected observatories fulfill the conditions required for stochastic analysis.

3. Stochastic processes

In the time domain, the series under study are considered as the specific result of a particular family of variables representing a statistical phenomenon that changes over time. Such family of variables is known as a stochastic process.

Any deductions from a single performance of the process require the imposition of restrictions: To be stationary in mean and variance and ergodic. Moreover, the sample must have an approximated normal distribution in order to satisfy the application rules in the autocorrelations of the later analysis. Stochastic processes can be represented as linear combinations of terms in which random variables take part, and, in particular, those formed by random pure processes (white noise), autoregressive processes, and moving average processes.

An autoregressive process of order p, denominated AR(p), is that in which a term of the series X_t can be expressed as a linear combination of the preceding p terms plus one at random

$$X_t = f_1 X_{t-1} + f_2 X_{t-2} + \ldots + f_{t-p} X_{t-p} + e_t$$

so that, using the linear operator of lag $L X_t = X_{t-1}$ gives

$$f(L) X_t = e_t.$$

A moving average process of order q, denominated MA(q), is that in which a term of the series X_t can be expressed as a linear combination of the preceding random q terms

$$X_t = e_t + q_1 e_{t-1} + q_2 e_{t-2} + \ldots + q_q e_{t-q}.$$

An ARMA (p, q) process is autoregressive, of moving average, from the combination of the former.

$$X_t = f_1 X_{t-1} + \ldots + f_{t-p} X_{t-p} + e_t + q_1 e_{t-1} + \ldots + q_q e_{t-p}$$

The reason for devising an ARIMA model is to find a function that, with a limited number of parameters, is a sufficiently close representation of how the observed series behaves.

The key mechanism for identifying a stochastic process is the calculation of the autocorrelation function (ACF) and partial autocorrelation function (PACF), that consider the values of the time series in relation with its lags, thereby revealing its inner structure. The values of the mean m and sampling covariances g_k establish the properties and relationships of the autocorrelation coefficients r_k for the corresponding processes AR(p), MA(q), and ARMA(p, q).

The autocorrelation coefficient, r_k, where $k = 0, 1, \ldots$, indicates the lag interval between the terms of the series. The sequence of r_k values constitutes the estimated autocorrelation function (or correlogram) calculated from the values of the series. When r_k is used as estimator of the stochastic processes, its distribution must be known in order to make inferences and to verify the significance of the different coefficients.

The partial autocorrelation coefficient (PACF) of order k is defined as the measurement of the linear relationship between k separate lags, independently of the intermediate values.

It is possible to demonstrate theoretically that in autoregressive AR(p) processes, a system of equations is obtained relating the first p coefficients of autocorrelation with the parameters ϕ of the process (Yule-Waker equations). In autoregressive processes, the autocorrelation function decreases after a determined lag, geometrically or sinusoidally with possible alternation of sign, but without becoming zero. This makes it difficult to interpret the correlogram for establishing the order: It is necessary to use the function of partial autocorrelation coefficient. The Yule-Waker equations enable the coefficients ϕ_i to be determined. The partial autocorrelation coefficients will be different from zero for lags not exceeding p, as by hypothesis the coefficients $\phi_1, \phi_2, \ldots, \phi_p$ are non-zero.

In the moving average MA(q) processes, it is possible to obtain the relationship of the autoregression coefficients with the θ parameters of the model. Autocorrelation coefficients r_k are zero for lags above q, and the partial autocorrelation coefficients decrease exponentially and/or sinusoidally.

Thus, the behavior of the autocorrelation function of a MA process seems to be similar to that of the partial autocorrelation function of an AR process.

In the ARMA(p, q) processes, the identification of parameters is more complex because, generally, in the ACF or in the PACF the first values do not behave constantly and the following ones present damped oscillation, with no characteristic pattern. Although identifying mechanisms have been proposed, the most used system is that of comparison with the theoretical behavior of pre-established generative functions for values below p and q (Uriel, 1995).

The ARIMA (p, d, q) processes are ARMA (p, q) processes related to series that, not being initially stationary, have been subjected to an order-d differencing process. That is, an "integrated" process is denominated ARIMA (p, d, q) if, taking differences of order d, is transformed into an ARMA (p, q) −type stationary process.

Once the generative model is identified, the corresponding values of the parameters are calculated from the known correlation coefficients.

Fit is verified by adaptation of the model, using analysis of residuals -the difference between observed and obtained values. Ideal models are those of best fit with lowest number of parameters. There are two methods of verification: by autocorrelation of residuals, or by residual variance. Autocorrelation of residuals is based on their constituting white noise and random time distribution. Consequently, subjected to correlational analysis, all the coefficients of the correlogram of residuals must be within the confidence interval, being significantly zero. The alternative analysis of residual variance establishes the best order of the autoregressive model to avoid overfit. This addresses the significance of the extra variance explained by the introduction of a new parameter.

4. Application to the series of temperatures. Results

The presentation of the procedure and results follows the sequence of phases presented above.

4.1 Determination of order

We have stated that the order of the p autoregressive coefficients and of the q moving averages in each series is determined using the autocorrelation coefficients (ACF) and partial autocorrelation coefficients (PACF). Briefly, if the coefficients in the first q lags of the ACF are non-zero and the following ones are zero, and those of the PACF decrease exponentially or sinusoidally, we initially ascribe the series as MA(q). On the other hand, if the coefficients in the first p lags of the PACF are non-zero and the following ones are zero, and those of the ACF decrease exponentially or sinusoidally, we initially ascribe the series as AR(p). If the form of the ACF and PACF is a superimposition of those described, the model is denominated ARMA(p, q).

To avoid reiteration in the graphical display of the procedure and observations, whose results are implicit in the following tables, we will give later, as significant examples, only the monthly series of September and March of the minimum temperatures from the Sevilla observatory (Figures 2, 3, 4, 5).

As the analysis is referred to monthly and annual series, which do not present stationarity, we determine the orders of the ARIMA(p, d, q) models. Because the series present a trend, they are subjected to first-order differencing so that, in such cases, $d = 1$.

The graphs (Figures 2-5) display the simple autocorrelation and partial autocorrelation functions of the minimum temperatures for the months of September and March at Sevilla. The first conclusion suggests both series belong to an MA(1) model, since only the first lag of the ACF is significantly different from zero (although one of the later retards is very close to the level of confidence). The shape of the first lags of the ACF is descending. However, calculation of the corresponding coefficients shows that the case of September leads to a state at the limit of invertibility (the theoretical condition for rejecting the model). Thus, we reformulate the model as AR(3), since the three autoregressive coefficients prove to be significant. Given that the series are subject to a trend, we introduce first-

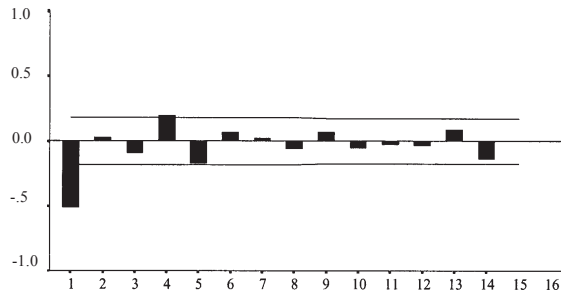Fig. 2. Autocorrelation function of the series of minimum temperatures in September at Sevilla.
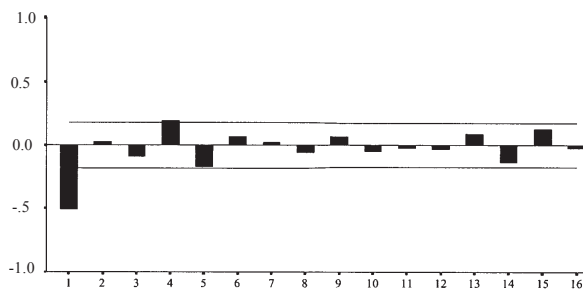

Fig. 3. Partial autocorrelation function of the series of minimum temperatures in September at Sevilla.
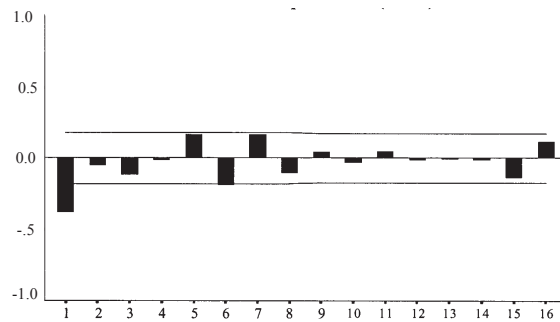

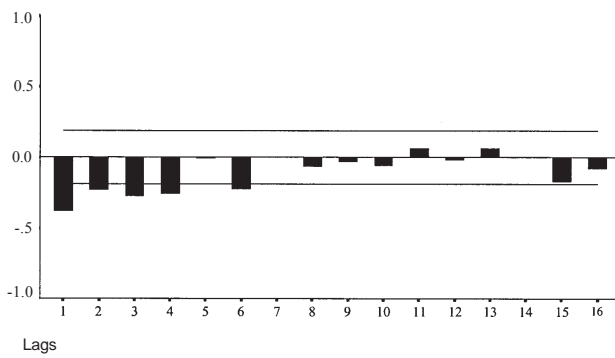Fig. 4. Autocorrelation function of the series of minimum temperatures in March at Sevilla.

Fig. 5. Partial autocorrelation function of the series of minimum temperatures in March at Sevilla.

Lags

order differencing and conditionally accept models of type (3,1,0) for September and (0,1,1) for March. In both cases, logarithmic transformation is used to guarantee homocedasticity, so that the residuals are expressed on that scale. (This does not affect the internal correlation.) With this hypothesis, we construct the estimated series.

The models fulfilling the conditions *(p, d, q)* adequately represent the respective series of observations. Nevertheless, this does not rule out the existence of another model with different orders and a better fit. Before making predictions, it is useful to carry out an overfit such that among different models (also subjected to the whole procedure described) the criterion of quality is established by the lowest residual variance of the estimated series. (The residual variance is an estimator obtained from the squared sum of the residuals divided by the degrees of freedom of the series (n-2)).

Table 1 shows the orders *(p, d, q)* of the definitive models for each monthly series of maximum and minimum temperatures at the selected observatories.

We can see that the series of minimum temperatures with one order of derivation and logarithmic transformation fit pure models: 40% moving average −36% MA(1), although with only one case in Huelva− and the remaining 60% autoregressive models, mostly AR(2) (24 %).

The orders of maximum temperatures fit principal models of moving average MA(1), with Badajoz and Sevilla showing very similar values for the corresponding months. At Córdoba, autoregressive models predominate. For January at Sevilla, and April and annually at Huelva, orders of mixed autoregressive moving average AR(1) MA(1) models are produced, so that the first-derivation treatment yields ARIMA(1, 1, 1).

Table 1. Determination of the order (p, d, q) of the time series of maximum (left) and minimum (right) temperatures.

	Maximum temperatures				Minimum temperatures			
	Badajoz	Córdoba	Huelva	Sevilla	Badajoz	Córdoba	Huelva	Sevilla
Sep	(0, 1, 1)	(3, 1, 0)	(3, 1, 0)	(0, 1, 1)	(3, 1, 0)	[0, 1, 1]	(3, 1, 0)	(3, 1, 0)
Oct	(4, 1, 0)	(3, 1, 0)	(3, 1, 0)	(4, 1, 0)	(0, 1, 1)	(1, 1, 0)	(2, 1, 0)	(0, 1, 1)
Nov	(0, 1, 1)	(3, 1, 0)	(0, 1, 1)	(0, 1, 1)	(2, 1, 0)	(1, 1, 1)	(0, 1, 1)	(4, 1, 0)
Dec	(0, 1, 1)	(3, 1, 0)	(2, 1, 0)	(0, 1, 1)	(3, 1, 0)	(4, 1, 0)	(2, 1, 0)	(4, 1, 0)
Jan	(0, 1, 1)	(3, 1, 0)	(1, 1, 0)	(1, 1, 1)	(0, 1, 1)	(3, 1, 0)	(2, 1, 0)	(4, 1, 0)
Feb	(0, 1, 1)	(2, 1, 0)	(2, 1, 0)	(0, 1, 1)	(2, 1, 0)	(2, 1, 0)	(2, 1, 0)	(3, 1, 0)
Mar	(0, 1, 1)	(0, 1, 1)	(0, 1, 1)	(0, 1, 1)	(0, 1, 2)	(2, 1, 0)	(3, 1, 0)	(0, 1, 1)
Apr	(0, 1, 2)	(0, 1, 2)	(1, 1, 1)	(0, 1, 1)	(0, 1, 1)	(4, 1, 0)	(1, 1, 0)	(0, 1, 1)
May	(0, 1, 1)	(3, 1, 0)	(0, 1, 1)	(3, 1, 0)	(4, 1, 0)	(0, 1, 1)	(0, 1, 2)	(2, 1, 0)
Jun	(5, 1, 0)	(3, 1, 0)	(0, 1, 1)	(0, 1, 1)	(0, 1, 1)	(0, 1, 1)	(2, 1, 0)	(0, 1, 1)
Jul	[0, 1, 1]	(4, 1, 0)	(0, 1, 1)	(0, 1, 1)	[2, 1, 0]	(0, 1, 1)	(1, 1, 0)	[0, 1, 1]
Aug	(0, 1, 1)	(3, 1, 0)	(1, 1, 1)	(0, 1, 1)	(0, 1, 1)	(2, 1, 0)	(1, 1, 0)	(0, 1, 1)
Annual	(0, 1, 1)	(0, 1, 1)	(0, 1, 1)	[0, 1, 1]	(0, 1, 1)	(0, 1, 1)	(0, 1, 1)	[0, 1, 1]

We should point out that in repeated cases we have rejected MA(1) models due to their autocorrelations, but in which the parameter q_1, probably well estimated, reaches values close to unity ($\theta_1 = 0.999\ldots$), situating the model at the limit of invertibility. In these cases, we preferred to look for the complementary autoregressive model, often of order 3 and above, even though causing the residual variance to increase. (There is a dual relationship, so that theoretically an MA(1) process is equivalent to an AR(∞) one with coefficients decreasing in geometric progression; in practice, we looked for the complementary AR(n) that satisfies the general requirements.)

The values given in brackets [ ] in Table 1, indicate that their acceptance is subject to the possible heterocedasticity of the series.

4.2 Identification of the model

The orders *(p, d, q)* established in the monthly series act as a basis for calculating the respective parameters ϕ_p, θ_q of the equation of the model in each series.

The initial acceptance of these values requires that statistically they are significantly non-zero at the level of confidence 0.95. With the coefficients accepted provisionally, the new series of values estimated by the model is generated; this yields the residual series by difference between the series of observed values and the respective estimated ones.

An autoregressive process of order *p* for the variable temperature $\{T_t\}$ is represented by the equation

$$T_t = \phi_1 \, T_{t-1} + \ldots + \phi_p \, T_{t-p} + \varepsilon_t + c$$

where *f* are the parameters whose acceptance has been previously validated. In the same way, a moving average process of order *q* is represented as

$$T_t = e_t - q_1 \, e_{t-1} - \ldots - q_q \, e_{t-q} + c$$

where $\{\varepsilon_t\}$ is a white noise process, with e_t independent of the T_{t-i} values, and *c* is a constant.

Identification of the ARIMA model that represents the organized component of each series, besides being necessary before predictions can be made, is intrinsically valuable for establishing the time organization of the set of observations and for characterizing the internal structure of each series (Figures 6 and 7). The order of the parameters *(p, d, q)* of the model indicates the relationships between consecutive elements of the meteorological variable, and knowing it enables comparative classifications. As a mathematical model, autoregressive moving average functions constitute a synthetic expression of the complex physical phenomenon which they represent in a simplified way.

Tables 2, 3, 4 and 5 include all the parameters calculated for the monthly and annual models for
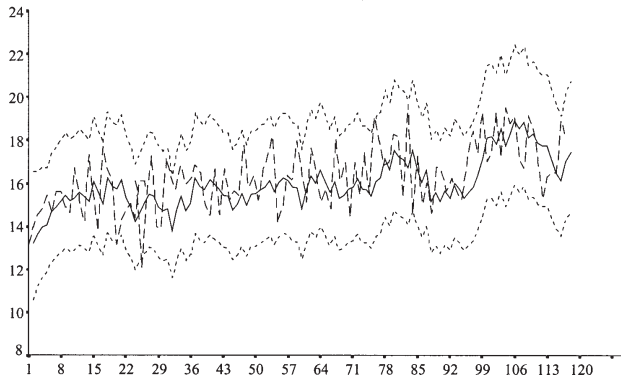
Fig. 6. Representation of the estimated model, with its confidence interval, superimposed on the series of observations of minimum temperature in September at Sevilla.
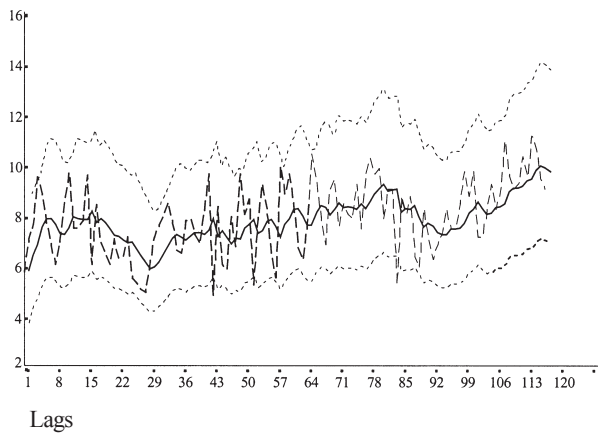
Fig. 7. Representation of the estimated model, with its confidence interval, superimposed on the series of observations of minimum temperature in March at Sevilla.

the observatories of the study. They enable us to establish the corresponding series estimated from the various models, and thereby obtain the residual series.

Each serial model *(p, d, q)* is validated from the non-correlation of its residual series. These must be random −the white noise process− with zero mean, constant variance (conditions fulfilled by the procedure applied if the series of observations is homocedastic), and non-correlated. For this, we see from the graphs of the simple autocorrelation function that all the lags are in fact within the interval of confidence. Thus, at the level of 0.95, the accepted residual correlation is significantly zero (Figs. 8 and 9).

5. Prediction

The first phases in devising the ARIMA model constitute an iterative process to formulate a model

Table 2. Identification of the parameters of the model of maximum and minimum temperature at Badajoz.

	Maximum temperatures Parameters of the model					Residual v.
Sep	$\theta_1=0.9026$					0.004
Oct	$\phi_1=-0.7265$	$\phi_2=-0.6111$	$\phi_3=-0.4796$	$\phi_4=-0.1932$		0.008
Nov	$\theta_1=0.8025$					0.007
Dec	$\theta_1=0.7995$					0.007
Jan	$\theta_1=0.9151$					0.008
Feb	$\theta_1=0.8862$					0.012
Mar	$\theta_1=0.8425$					0.012
Apr	$\theta_1=1.0491$	$\theta_2=0.2922$				0.009
May	$\theta_1=0.9264$					0.008
Jun	$\phi_1=-0.8784$	$\phi_2=-0.8403$	$\phi_3=-0.6110$	$\phi_4=-0.458$	$\phi_5=-0.2$	0.005
Jul	$\theta_1=0.8188$					0.002
Aug	$\theta_1=0.9010$					0.002
Annual	$\theta_1=0.6478$					0.001

	Minimum temperatures				Residual v.
Sep	$\phi_1=-0.8016$	$\phi_2=-0.5554$	$\phi_3=-0.3369$		0.007
Oct	$\theta_1=0.9102$				0.015
Nov	$\phi_1=-0.6462$	$\phi_2=-0.2095$			0.007
Dec	$\phi_1=-0.7906$	$\phi_2=-0.5545$	$\phi_3=-0.2268$		0.342
Jan	$\theta_1=0.9475$				0.149
Feb	$\phi_1=-0.5785$	$\phi_2=-0.2066$			0.416
Mar	$\theta_1=0.6313$	$\theta_2=0.2980$	$c=0.0029$		0.035
Apr	$\theta_1=0.8268$				0.014
May	$\phi_1=-0.3610$	$\phi_2=-0.7787$	$\phi_3=-0.4289$	$\phi_4=-0.2019$	0.010
Jun	$\theta_1=0.8776$				0.006
Jul	$\phi_1=-0.8044$	$\phi_2=-0.2350$			0.005
Aug	$\theta_1=0.9254$				0.004
Annual	$\theta_1=0.6601$				0.003

compatible with the structure of the data observed each month. The aim is to discover, within a level of confidence and from the knowledge of a period of effective recording, the subsequent behaviour of the series. The model devised is thus used in estimating the future values of the variables under study. The equations of the model, fitted by the elements of the series for $1 \le t \le n$, are used to progressively calculate the results of the variable t for $n + 1, n + 2, \ldots n + i$.

We fixed the end of the prediction period as the year 2010 for procedural reasons and because it coincides with proposals from other bodies. Increasing the prediction period means that it is

Table 3. Identification of the parameters of the model of maximum and minimum temperatures at Córdoba.

	Maximum temperatures				
	Parameters of the model				Residual v.
Sep	$\phi_1 = -0.9015$	$\phi_2 = -0.7253$	$\phi_3 = -0.2982$		0.004
Oct	$\phi_1 = -0.6591$	$\phi_2 = -0.6150$	$\phi_3 = -0.3320$		0.009
Nov	$\phi_1 = -0.6340$	$\phi_2 = -0.4277$	$\phi_3 = -0.3358$		0.009
Dec	$\phi_1 = -0.8871$	$\phi_2 = -0.6272$	$\phi_3 = -0.2381$		0.007
Jan	$\phi_1 = -0.6653$	$\phi_2 = -0.4437$	$\phi_3 = -0.3258$		0.009
Feb	$\phi_1 = -0.6785$	$\phi_2 = -0.4352$			0.012
Mar	$\theta_1 = 0.8867$				0.013
Apr	$\theta_1 = 0.8355$				0.009
May	$\phi_1 = -0.6738$	$\phi_2 = -0.5025$	$\phi_3 = -0.2961$		0.010
Jun	$\phi_1 = -0.7011$	$\phi_2 = -0.5471$	$\phi_3 = -0.3113$		0.006
Jul	$\phi_1 = -0.7373$	$\phi_2 = -0.5764$	$\phi_3 = -0.4894$	$\phi_4 = -0.2402$	0.002
Aug	$\phi_1 = -0.6402$	$\phi_2 = -0.6682$	$\phi_3 = -0.3812$		0.001
Annual	$\theta_1 = 0.7882$				0.001
	Minimum temperatures				Residual v.
Sep	$\theta_1 = 0.8480$				0.053
Oct	$\phi_1 = -0.4790$				0.018
Nov	$\phi_1 = -0.4910$				0.072
Dec	$\phi_1 = -0.8740$	$\phi_2 = -0.6582$	$\phi_3 = -0.5442$	$\phi_4 = -0.2487$	0.240
Jan	$\phi_1 = -0.6883$	$\phi_2 = -0.4409$	$\phi_3 = -0.2587$		0.370
Feb	$\phi_1 = -0.7481$	$\phi_2 = -0.3774$			0.207
Mar	$\phi_1 = -0.8654$	$\phi_2 = -0.7662$	$\phi_3 = -0.5712$	$\phi_4 = -0.3540$	0.279
Apr	$\theta_1 = 0.8230$				0.014
May	$\theta_1 = 0.8880$				0.008
Jun	$\theta_1 = 0.8661$				0.073
Jul	$\theta_1 = 0.8268$				0.006
Aug	$\phi_1 = -0.6106$	$\phi_2 = -0.3074$			0.044
Annual	$\theta_1 = 0.7948$				0.029

determined by the autoregressive elements, and the statistical limits of confidence are increased such that they are no longer of any practical use. Groups of experts on climate change have chosen this date for verifying control measures, and formulating different-angled hypotheses about change in climate variables.It seems advisable to coincide with this period of analysis –rather long but without loss of reliability in the estimations.

Evidently, the statistical model devised does not introduce external actions into the generative mechanism, nor does it assume actions entering from outside the system itself. However, we have

Table 4. Identification of the parameters of the model of maximum and minimum temperatures at Huelva.

	Maximum temperatures			Residual v.
	Parameters of the model			
Sep	$\phi_1 = -0.7714$	$\phi_2 = -0.4651$	$\phi_3 = -0.2725$	0.007
Oct	$\phi_1 = -0.7045$	$\phi_2 = -0.4615$		0.001
Nov	$\theta_1 = 0.8034$			0.002
Dec	$\phi_1 = -0.6894$	$\phi_2 = -0.3191$		0.012
Jan	$\phi_1 = -0.6552$	$\phi_2 = -0.3489$		0.012
Feb	$\phi_1 = -0.6035$	$\phi_2 = -0.3541$		0.009
Mar	$\phi_1 = -0.6415$	$\phi_2 = -0.4488$	$\phi_3 = -0.2587$	0.019
Apr	$\phi_1 = -0.5868$			0.001
May	$\theta_1 = 0.6296$	$\theta_2 = 0.2814$		0.007
Jun	$\phi_1 = -0.2689$	$\phi_2 = -0.2896$		0.007
Jul	$\phi_1 = -0.5018$			0.004
Aug	$\theta_1 = 0.8583$			0.002
Annual	$\phi_1 = -0.3252$	$\theta_1 = 0.9329$		0.002
	Minimum temperatures			Residual v.
Sep	$\phi_1 = -0.7714$	$\phi_2 = -0.4651$	$\phi_3 = -0.2725$	0.075
Oct	$\phi_1 = -0.7045$	$\phi_2 = -0.4615$		0.016
Nov	$\theta_1 = 0.8434$			0.024
Dec	$\phi_1 = -0.6894$	$\phi_2 = -0.3191$		0.124
Jan	$\phi_1 = -0.6552$	$\phi_2 = -0.3489$		0.120
Feb	$\phi_1 = -0.6035$	$\phi_2 = -0.3489$		0.090
Mar	$\phi_1 = -0.6415$	$\phi_2 = -0.4488$	$\phi_3 = -0.2587$	0.197
Apr	$\phi_1 = -0.5868$			0.010
May	$\theta_1 = 0.6296$	$\theta_2 = 0.2814$		0.078
Jun	$\phi_1 = -0.6289$	$\phi_2 = -0.2856$		0.074
Jul	$\phi_1 = -0.5018$			0.039
Aug	$\phi_1 = \quad \phi_2 = \quad \phi_3 = \quad \phi_4 = \quad \theta_1 = \quad \theta_2 = \quad c =$			0.115
Annual	$\theta_1 = 0.3035$			0.025

to stress that the possible effects of anthropogenic activity, and its constant intensification, are assimilated by the statistical model. Climate change has been progressively incorporated throughout the period of recordings and has affected the meteorological variables under study in this research. The formulated prediction implies that the actions on the climate system will continue in the near future, similar in intensity and progression to those of the recent past.

As a result of the methodological process described above (Tables 6, 7, and 8) the monthly predictions of maximum and minimum temperatures at the observatories of Badajoz, Sevilla, and

Table 5. Identification of the parameters of the model of maximum and minimum temperature at Sevilla.

	Maximum temperatures Parameters of the model				Residual v.
Sep	$\theta_1=0.8815$				0.003
Oct	$\phi_1=-0.7265$	$\phi_2=-0.6111$	$\phi_3=-0.4796$	$\phi_4=-0.1932$	0.006
Nov	$\theta_1=0.7628$				0.005
Dec	$\theta_1=0.8423$				0.003
Jan	$\phi_1=-0.2361$	$\theta_1=0.8869$			0.004
Feb	$\theta_1=0.8744$				0.007
Mar	$\theta_1=0.8342$				0.006
Apr	$\theta_1=0.8621$				0.007
May	$\phi_1=-0.7227$	$\phi_2=-0.6100$	$\phi_3=-0.4167$	$\phi_4=-0.1981$	0.008
Jun	$\theta_1=0.9248$				0.004
Jul	$\theta_1=0.9106$				0.001
Aug	$\theta_1=0.8983$	$c=0.0009$			0.002
Annual	$\theta_1=0.7701$				0.001
	Minimum temperatures				Residual v.
Sep	$\phi_1=-0.7773$	$\phi_2=-0.5623$	$\phi_3=-0.3612$		0.007
Oct	$\theta_1=0.7750$				0.011
Nov	$\phi_1=-0.8304$	$\phi_2=-0.6001$	$\phi_3=-0.5468$	$\phi_4=-0.3138$	0.042
Dec	$\phi_1=-0.8618$	$\phi_2=-0.6517$	$\phi_3=-0.4408$	$\phi_4=-0.2548$	0.166
Jan	$\phi_1=-0.7872$	$\phi_2=-0.6021$	$\phi_3=-0.4925$	$\phi_4=-0.1995$	0.182
Feb	$\phi_1=-0.7393$	$\phi_2=-0.4545$	$\phi_3=-0.2140$		0.167
Mar	$\theta_1=0.8029$				0.030
Apr	$\theta_1=0.7690$				0.012
May	$\phi_1=-0.5170$	$\phi_2=-0.4221$			0.009
Jun	$\theta_1=0.8340$				0.005
Jul	$\theta_1=0.7711$				0.002
Aug	$\theta_1=0.8833$				0.003
Annual	$\theta_1=0.8234$				0.004

Córdoba until the year 2010, from ARIMA models, are included. We should point out that expression of the annual predictions in hundredths of ºC is only of mathematical interest for observing variations between years; from the point of view of climatic significance, the precision for temperature is of tenths of a degree, as in the records of the series.

In order to compare the change obtained from preceding studies on the linear trend of the set of historical series with the future predictions from the devised models, we calculated the seasonal and annual values from the respective monthly estimations. Tables 9 and 10 present the trend of the series of historical data and the future estimations.
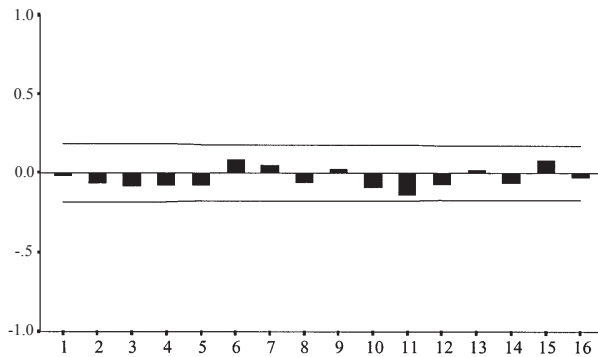
Fig. 8. Verification, using the correlograms of the residual series, of the fit of the devised model for the series of maximum and minimum temperatures in September at Sevilla.
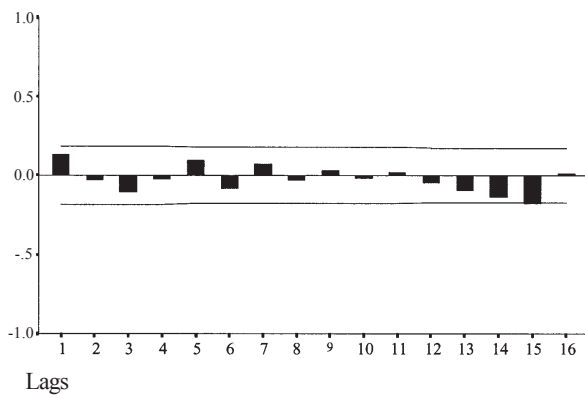
Fig. 9. Verification, using the correlograms of the residual series, of the fit of the devised model for the series of maximum and minimum temperatures in March at Sevilla.

In the estimation of maximum temperatures, the warming in spring, common to all the observatories −and very intense at Badajoz− is striking, when, historically, they have remained stable. There is a marked divergence in thermal behavior predicted for the summers: A rise at Córdoba and fall at Sevilla, reflected in the annual warming at Córdoba against the cooling at Sevilla. The lack of statistical significance for the historical trend, and its inconstancy from one observatory to another, perhaps lead to predictions that are not concordant.

The predictions for minimum temperatures at Badajoz are similar to those at Sevilla (perhaps as a result of the relative similarity of the monthly orders of the respective *p, d, q* models). At Córdoba, the magnitude of the predicted decrease in winter minimum temperatures is surprising, and moderates the annual general warming. In contrast, both Sevilla and Badajoz show an even greater intensification of the rising trend for the near future, in accord with the results obtained by extrapolation of the polynomial line of the historical series.

Table 6. Estimation of the maximum and minimum temperatures (°C) predicted by the model for Badajoz.

	S	O	N	D	J	F	M	A	M	J	J	A
					Maximum temperatures							
2001	29.84	24.87	19.60	14.56	13.79	16.09	20.95	22.61	25.11	30.80	34.70	34.04
2002	29.83	24.78	19.60	14.57	13.79	16.09	20.98	22.64	25.11	30.04	34.70	34.03
2003	29.83	24.85	19.61	14.58	13.79	16.08	21.00	22.67	25.11	31.07	34.71	34.02
2004	29.82	25.03	19.62	14.58	13.79	16.08	21.03	22.70	25.10	31.19	34.71	34.01
2005	29.82	25.00	19.62	14.59	13.79	16.08	21.05	22.73	25.10	30.71	34.71	34.00
2006	29.81	24.93	19.63	14.60	13.79	16.07	21.07	22.76	25.10	30.79	34.72	33.99
2007	29.81	24.93	19.63	14.60	13.79	16.07	21.10	22.79	25.10	30.82	34.72	33.98
2008	29.80	24.99	19.64	14.61	13.79	16.06	21.12	22.82	25.09	30.74	34.72	33.96
2009	29.80	25.02	19.65	14.62	13.79	16.06	21.15	22.85	25.09	30.97	34.73	33.95
2010	29.79	25.01	19.65	14.62	13.79	16.06	21.17	22.88	25.09	30.95	34.73	33.94
					Minimum temperatures							
	S	O	N	D	J	F	M	A	M	J	J	A
2001	17.76	13.00	8.78	3.66	5.14	4.19	7.85	10.20	14.00	16.16	19.41	18.79
2002	17.45	13.03	8.94	2.95	5.18	4.41	7.87	10.23	14.12	16.69	19.46	18.81
2003	17.34	13.07	8.89	3.81	5.23	4.31	7.89	10.26	14.14	16.72	19.49	18.84
2004	17.50	13.10	8.93	3.64	5.27	4.31	7.92	10.29	14.17	16.75	19.52	18.86
2005	17.63	13.13	8.96	3.56	5.32	4.32	7.94	10.32	14.19	16.78	19.56	18.88
2006	17.57	13.16	8.97	3.52	5.37	4.30	7.96	10.35	14.25	16.81	19.59	18.90
2007	17.59	13.20	9.00	3.68	5.42	4.30	7.99	10.38	14.30	16.84	19.63	18.92
2008	17.66	13.23	9.02	3.65	5.46	4.29	8.01	10.40	14.33	16.87	19.66	18.95
2009	17.70	13.26	9.04	3.65	5.51	4.29	8.03	10.43	14.37	16.90	19.70	18.97
2010	17.71	13.30	9.07	3.67	5.56	4.28	8.06	10.46	14.41	16.93	19.73	18.99

6. Conclusions

Autoregressive Integrated Moving Average (ARIMA) models have been devised for the long-term monthly series of maximum and minimum temperatures from observatories in south-west Spain. Previously, it was confirmed that the first derivation and logarithmic transformation stabilized the mean and variance of the original series, and converted them into stationary series.

The orders (p, d, q) for each monthly series have been obtained. The series of minimum temperatures with one order of derivation and logarithmic transformation have been fitted to pure moving average (MA) models (38%) –above all MA(1)– to autoregressive (AR) models (58%) –mostly AR(2)– and to some mixed (ARMA) models (4%). The maximum temperatures fit principal moving average models MA(1), with very similar monthly values for Badajoz and Sevilla, whereas at the Córdoba observatory, the predominating models are autoregressive, AR (3).

Table 7. Estimation of the maximum and minimum temperatures (°C) predicted by the model for Sevilla.

	S	O	N	D	J	F	M	A	M	J	J	A
					Maximum temperatures							
2001	31.21	27.00	20.70	16.71	16.10	18.27	22.74	23.99	26.76	31.04	35.11	34.60
2002	31.19	27.02	20.68	16.70	16.08	18.26	22.76	23.99	27.00	31.03	35.10	34.57
2003	31.17	27.02	20.66	16.70	16.07	18.25	22.77	24.00	27.40	31.01	35.09	34.53
2004	31.15	27.09	20.64	16.69	16.06	18.24	22.79	24.00	27.10	30.99	35.07	34.50
2005	31.14	27.06	20.62	16.69	16.05	18.23	22.81	24.01	26.94	30.97	35.06	34.47
2006	31.12	27.00	20.60	16.68	16.04	18.22	22.83	24.02	27.01	30.96	35.04	34.44
2007	31.10	26.99	20.58	16.68	16.03	18.21	22.85	24.02	27.09	30.94	35.03	34.41
2008	31.08	27.00	20.56	16.67	16.02	18.20	22.86	24.03	27.10	30.92	35.02	34.37
2009	31.06	27.00	20.54	16.67	16.01	18.19	22.88	24.03	27.04	30.90	35.00	34.34
2010	31.05	26.98	20.52	16.66	16.00	18.18	22.90	24.04	27.02	30.89	34.99	34.31
					Minimum temperatures							
	S	O	N	D	J	F	M	A	M	J	J	A
2001	18.23	14.10	10.77	7.31	6.74	6.73	9.75	11.29	14.49	17.35	19.76	19.45
2002	17.99	14.13	11.01	7.13	6.24	6.29	9.78	11.31	14.50	17.37	19.79	19.47
2003	17.87	14.15	10.78	6.69	6.30	6.72	9.80	11.33	14.47	17.40	19.82	19.49
2004	17.98	14.17	11.23	7.48	6.69	6.66	9.83	11.36	14.54	17.42	19.85	19.51
2005	18.16	14.19	11.07	7.41	6.73	6.62	9.85	11.38	14.58	17.45	19.88	19.54
2006	18.11	14.22	11.11	7.34	6.61	6.60	9.88	11.41	14.59	17.47	19.91	19.56
2007	18.11	14.24	11.13	7.35	6.56	6.66	9.90	11.43	14.63	17.50	19.95	19.58
2008	18.18	14.26	11.16	7.36	6.66	6.64	9.93	11.46	14.67	17.52	19.98	19.60
2009	18.25	14.28	11.28	7.55	6.75	6.64	9.95	11.48	14.69	17.55	20.01	19.62
2010	18.26	14.31	11.27	7.54	6.75	6.65	9.98	11.51	14.73	17.58	20.04	19.64

The parameters θ_p, ϕ_q of the estimated functions of the monthly stochastic models have also been obtained, and subjected to validation processes to test whether the corresponding series of residuals are random. These values, and the residual variance, are given in the respective tables.

From the estimated functions for each month, predictions of change have been formulated: Those from the ARIMA models of minimum temperature suggest that the study region is in a phase of annual mean thermal increase of the order of 0.2 °C in the next decade. However, the change in maximum temperatures is not significantly defined. These estimations are in agreement with those obtained in earlier studies using other methods, but because of the better goodness-of-fit, the level of precision is much higher in this case.

Table 8. Estimation of the maximum and minimum temperatures predicted by the model for Córdoba.

	S	O	N	D	J	F	M	A	M	J	J	A
						Maximum temperatures						
2001	30.86	26.14	18.75	15.18	14.84	17.97	21.91	23.40	26.45	32.66	35.64	36.14
2002	30.19	25.98	19.27	15.28	14.99	18.44	21.94	23.41	27.20	32.95	36.10	35.85
2003	30.19	26.24	19.25	15.09	14.93	18.31	21.97	23.41	26.78	32.58	36.07	35.27
2004	30.49	26.10	19.19	15.16	14.93	18.22	22.01	23.41	26.73	32.43	36.01	35.45
2005	30.39	26.15	19.08	15.24	14.94	18.37	22.04	23.41	26.78	32.68	35.85	35.82
2006	30.23	26.14	19.20	15.22	14.99	18.34	22.07	23.41	26.92	32.74	35.89	35.67
2007	30.33	26.20	19.21	15.22	14.99	18.33	22.11	23.42	26.84	32.65	35.98	35.44
2008	30.36	26.18	19.21	15.26	15.00	18.38	22.14	23.42	26.84	32.64	35.97	35.54
2009	30.28	26.20	19.19	15.27	15.01	18.38	22.17	23.42	26.86	32.72	35.93	35.68
2010	30.39	26.21	19.32	15.23	14.96	18.35	22.19	23.43	26.84	32.74	35.95	35.72
						Minimum temperatures						
	S	O	N	D	J	F	M	A	M	J	J	A
2001	18.09	12.87	9.90	7.92	7.66	6.86	9.37	11.12	14.39	17.93	20.72	21.19
2002	18.11	13.40	10.22	8.41	7.12	5.53	9.15	11.14	14.42	17.95	20.73	21.12
2003	18.12	13.14	10.10	7.42	6.58	6.67	9.13	11.16	14.44	17.98	20.75	21.17
2004	18.13	13.28	10.19	6.47	6.55	6.39	9.00	11.18	14.47	18.00	20.77	21.22
2005	18.15	13.22	10.18	7.54	7.08	6.14	9.13	11.20	14.50	18.02	20.78	21.23
2006	18.16	13.25	10.21	7.63	7.13	6.44	9.27	11.22	14.53	18.05	20.80	21.26
2007	18.17	13.24	10.22	7.83	6.98	6.33	9.21	11.24	14.55	18.07	20.81	21.29
2008	18.19	13.25	10.25	7.45	6.95	6.30	9.21	11.26	14.58	18.09	20.83	21.32
2009	18.20	13.25	10.27	7.38	7.05	6.38	9.21	11.28	14.61	18.11	20.85	21.35
2010	18.21	13.26	10.29	7.61	7.15	6.35	9.27	11.30	14.63	18.14	20.86	21.38

Table 9. Comparison of the seasonal change in maximum temperature of the historical series with the predictions of the model for the next decade.

	Autumn	Winter	Spring	Summer	Annual
		Maximum temperatures. Linear trend			
		Instrumental series			
Badajoz	-0.01	-0.01	0.01	-0.09	-0.02
Córdoba	-0.01	-0.07	0.01	0.11	-0.02
Sevilla	-0.09	-0.07	-0.02	-0.02	-0.01
		Prediction			
Badajoz	-0.01	0.01	0.17	0.04	0.05
Córdoba	0.09	0.03	0.11	0.21	0.18
Sevilla	-0.19	-0.11	0.05	-0.21	-0.11

Table 10. Comparison of the seasonal change in minimum temperatures of the historical series with the predictions of the model for the next decade.

	Autumn	Winter	Spring	Summer	Annual
	Minimum temperatures. Linear trend				
	Instrumental series				
Badajoz	0.24	0.23	0.28	0.21	0.23
Córdoba	0.18	0.27	0.18	0.10	0.18
Sevilla	0.19	0.24	0.19	0.21	0.16
	Prediction				
Badajoz	0.27	0.20	0.31	-0.10	0.28
Córdoba	0.00	-0.23	0.22	0.23	0.06
Sevilla	0.35	0.03	0.28	0.26	0.20

References

Al-awadhi S. and J., Jolliffe, 1998. Times series modelling of surface presure data. *Int. J. Climatol.* **18,** 443-455.

Bravo, J. L., M. M. Nava and C. Gay, 2001. Linear and regressive stochastic models for prediction models of daily maximum ozones values at Mexico City atmosphere. *Atmósfera* **14**, 113-123.

Box, G.E.P. and G. M. Jenkins, 1970. *Time Series Analysis, Forecasting and Control.* Holden-Day. San Francisco.

Chatfield, C., 1992. *The analysis of time series*. Chapman & Hall. London.

Ferrand, M., 1997. *Programación y análisis estadístico.* MacGraw-Hill. Madrid.

García-Barrón, L., 2000. *Análisis de series termopluviométricas para la elaboración de modelos climáticos.* Universidad de Sevilla.

García-Barrón, L. and M. F. Pita, 2001. Propuesta metodológica para la determinación de inhomogeneidades relativas en las series de observaciones. *El tiempo del clima.* Publicaciones de la Asociación Española de Climatología, serie A Vol. **2**, 87-96.

García-Barrón, L., 2002. Un modèle pour l'analyse de la sécheresse dans les climats mediterranéens. Publications de l´Association Internationale de Climatologie **14,** 67-73.

Garrido, J. and J. A. García, 1994. Aplicación de procesos autoregresivos-media móvil para modelizar series temporales de precipitación mensual en la España peninsular. *Anales de Física* **89**, 50-56.

Grieser. J., S. Trömel and C. Schönwiese, 2002. Statistical time series decomposition into significant components and application to European temperature. *Theor. Appl. Climatol.* **71**, 171-183.

Hurrell, J. M., 1995. Decenal trends in North Atlantic Oscillation and relationship to regional temperature precipitation. *Science* **269,** 676-679.

Kendall, S. M. and J. K. Ord, 1990. *Time series.* Edward Arnold, London.

Matyasovszky, I., 2002. A nonlinear approach to modeling climatological time series. *Theor. Appl. Climatol.* **69**, 139-147.

Martin M., L. V. Cremades and J. Santabarbara, 1999. Analysis and modelling of times series of surface wind speed and direction. *Int. J. Climatol.* **19**, 197-209.

Peña Sánchez de Rivera, D*.,* 1995. *Estadística, modelos y métodos*. Alianza Universitaria, Madrid.

Pita López, M. F., J. M. Camarillo and M. Aguilar, 1999. L´evolution de la variabilité pluviometrique en Andalousie (Espagne). Publications de l´Association Internationale de Climatologie **10**, 313-321.

Pozo-Vázquez, D., M. J. Esteban-Parra, F. S. Rodrigo and Y. Castro-Díez, 2000. An analysis of the variability of the North Atlantic Oscillation in the time and the frecuency domains. *Int. J. Climatol.* **20**, 1675-1692.

Solange, L. and J. Pinto Peixoto, 1996. The autoregresive model of climatological times series: an aplication to the longest times series in Portugal. *Int. J. Climatol.* **16**, 1165-1173.

Uriel, Z., 1995. *Análisis de series temporales.* Paraninfo, Madrid.

Wei, W. W., 1990. *Times series*. Addison-Wesley, Redwood City.

Zekai S., 1998. Small sample estimation of the variance of times-averages in climatic times series. *Int. J. Climatol.* **18**, 1725-1732.