Pergamon

PII: S0038-092X(98)00078-4

Solar Energy Vol. 63, No. 5, pp. 283–291, 1998

0038-092X/98 $—see front matter

MULTIPLICATIVE ARMA MODELS TO GENERATE HOURLY SERIES OF GLOBAL IRRADIATION

LL. MORA-LÓPEZ* and M. SIDRACH-DE-CARDONA‡**

* Dpto. Lenguajes y C. Computación. E.T.S.I. Informática. Universidad de Málaga, 29071 Málaga, Spain

** Dpto. Física Aplicada II. E.T.S.I. Informática. Universidad de Málaga, 29071 Málaga, Spain

Received 11 December 1997; revised version accepted 29 June 1998

Communicated by RICHARD PEREZ

Abstract—A methodology to generate hourly series of global irradiation is proposed. The only input parameter which is required is the monthly mean value of daily global irradiation, which is available for most locations. The procedure to obtain new series is based on the use of a multiplicative autoregressive moving-average statistical model for time series with regular and seasonal components. The multiplicative nature of this models enables us to capture the two types of relationships observed in recorded hourly series of global irradiation: on the one hand, the relationship between the value at one hour and the value at the previous hour; and on the other hand, the relationship between the value at one hour in one day and the value at the same hour in the previous day. In this paper the main drawback which arises when using these models to generate new series is solved: namely, the need for available recorded series in order to obtain the three parameters contained in the statistical ARMA model which is proposed (autoregressive coefficient, moving-average coefficient and variance of the error term). Specifically, expressions which enable estimatation of these parameters using only monthly mean values of daily global irradiation are proposed in this paper. © 1998 Elsevier Science Ltd. All rights reserved.

1. INTRODUCTION

Time series statistical models have been used to study the stationary and sequential characteristics of hourly series of global irradiation. The purpose of some of these studies is to analyze past data in order to infer statistical properties which capture the observed empirical regularities of the series and, then, use these properties to forecast its future behaviour.

This methodology has been used in the study of daily and hourly series of global irradiation (e.g. Brinkworth, 1977; Bendt *et al.*, 1981; Boch *et al.*, 1981; Aguiar *et al.*, 1988; Guinea *et al.*, 1988; Palomo, 1989; Aguiar and Collares-Pereira, 1992a).

In some of these works, not only are irradiation series statistically analyzed, but also methods which enable generation of synthetic series of irradiation are proposed. For instance, it is worth mentioning the work performed by Aguiar *et al.* (1988), where a simple method to generate daily series of global irradiation is proposed. This method is based on the use of first-order Markov matrices. The only parameter which is required to generate new series is the monthly mean value of the clearness index. The matrices which are used to generate new series prove to be universal.

The methodology of a "cascade of models" is the most widely used to generate hourly series of global irradiation and it is also the methodology which yields the best results. For example, Graham and Hollands (1990) use stochastic models to generate synthetic sets of hourly irradiation values, using daily values as input; Aguiar and Collares-Pereira (1992b) propose a time-dependent autoregressive Gaussian model (TAG): first, a daily series of global irradiation is generated and then hourly values are generated from these daily values.

A different methodology to analyze and generate hourly series of global irradiation was introduced in Mora-López and Sidrach-de-Cardona (1997): each monthly series of hourly values is directly analyzed and characterized and then new series with the same characteristic parameters are generated. In that work, a method to remove the stationary and daily trends which are observed in hourly irradiation series was proposed. The resultant monthly series are assumed to follow a seasonal multiplicative autoregressive moving-average (ARMA) model: a first-order autoregressive model ARMA(1,0) for the regular part and a first-order moving-average model ARMA(0,1) for

‡Author to whom correspondence should be addressed. ISES Member.

the seasonal part (see Box and Jenkins, 1976). The main problem with this methodology is that the underlying parameters of the model vary from one location to another and from one month to another. Therefore, as it had been pointed out in Mora-López and Sidrach-de-Cardona (1997), it is necessary to examine further the relationship between the parameters of the model and the available data from most locations.

The precise object of this paper is to find the relationship between the parameters of the underlying statistical model and the data which are usually available, and the relationship between different coefficients which correspond to different series. With this purpose, the relationship between the various coefficients in a model and some other values (such as monthly mean value of the daily clearness index, month of the year, etc.) will be examined.

2. DATA SET

The data of hourly series of global irradiation, $[G_h(t)]$, which are used in this work have been recorded throughout several years in various Spanish meteorological stations. These data are contained within the Data Base of Spanish Institute for Renewable Energy (CIEMAT). In Table 1 the number of available observations is reported, specifying year and latitude of the place where they were recorded (North latitude). Each year there are some months for which data are not available; for this reason, the total number of months for which observations are available in each location appears in the third column. This table also includes annual average values of daily global irradiatiorm, G_{dy}.

Table 1. Data set

Location	Years	Months	Lat. (°N)	G_{dy} (MJm^{-2})
Badajoz	1976–1983	84	38.89	16.6
Castellón	1979–1984	55	39.95	15.8
Madrid	1979–1986	96	40.45	16.6
Málaga	1977–1984	94	36.66	16.9
Murcia	1977–1984	96	38.00	16.9
Oviedo	1977–1984	96	43.35	11.1
Mallorca	1977–1984	94	39.33	15.5
Sevilla	1977–1984	78	37.42	17.6
Tortosa	1980–1984	52	40.81	15.1

3. STATISTICAL MODEL

The hourly series of global irradiation, $[G_h(t)]$ are not stationary. As it was described in Mora-López and Sidrach-de-Cardona (1997), stationary series can be obtained by removing the observed daily and seasonal trends with the following steps:

(1) Use the maximum hourly global irradiation, $G_{h,max}(t)$, to obtain:

$$X_h(t) = G_h(t)/G_{h,max}(t) \tag{1}$$

where $G_{h,max}(t)$ is computed using:

$$G_{h,max}(t) = 3.96[\cos\theta_z(t)]^{1.05}(MJm^{-2}) \tag{2}$$

where θ_z is the zenith angle. This is a clear-sky model estimated in Mora-López (1995), where the coefficients in the expression are computed using hourly series of global irradiation data recorded in various Spanish locations.

(2) Use the difference operator to obtain:

$$Y(t) = X_h(t) = X_h(t-s) \tag{3}$$

where s is the number of hours, per day, considered for each month—the value of s depends on the mean duration of a day in the corresponding month, and t ranges between s and the total hours of month. For the Spanish locations which have been used, this number is: $s=10$ hours for January and December; $s=12$ for February, October and November; $s=16$ for June and July; and $s=14$ for the remaining months. Using these values of s, those values which correspond to small fractions of an hour are removed.

For the new series obtained in this way, $Y(t)$, a seasonal multiplicative ARMA$(1,0)x(0,1)_s$ model is proposed, that is, a model whose regular part follows a first-order autoregressive model and whose seasonal part follows a first-order moving-average model:

$$Y(t) = \phi_1 Y(t-1) + \epsilon(t) - \theta_1\epsilon(t-s) \tag{4}$$

Using this model, the parameters ϕ_1: autoregressive coefficient, θ_1: moving average coefficient and σ^2: variance of the white noise process $\epsilon(t)$, have been estimated for each series. These parameters are estimated using maximum likelihood techniques (assuming a Gaussian distribution with zero mean for the white noise process). This method requires optimization and selects, among all possible values of the parameters, those for which the data set used for the estimation proves to be more likely.

The procedure which is proposed in Mora-López and Sidrach-de-Cardona (1997) to generate new hourly exposure series of global irradiation consists of several steps, which are described below, and requires knowledge of the parameters ϕ_1, θ_1 and σ^2 and the mean value of the series $[X_h(t)]$ which appears in eqn (1). However, we have verified that it is possible to use a similar procedure to this one which does not require knowing the mean value of the series $[X_h(t)]$.

In order to use the proposed ARMA model, in the next sections the relationships which have been detected between the parameters of the model and other parameters which characterize hourly series of global irradiation are described.

4. CHARACTERIZING THE PARAMETERS IN THE ARMA MODEL

First of all, it has been analyzed whether there is any relatiopship between the three estimated parameters in the proposed ARMA model, namely: the autoregressive coefficient $\hat{\phi}_1$, the moving-average coefficient $\hat{\theta}_1$ and the variance of the white noise process, $\hat{\sigma}^2$. In Figs. 1–3 various two-dimensional plots have been depicted: $\hat{\phi}_1$ vs. $\hat{\theta}_1$, $\hat{\phi}_1$ vs. $\hat{\sigma}^2$, and $\hat{\theta}_1$ vs. $\hat{\sigma}^2$, respectively, for all the data used. In these figures it is observed that there does not seem to be a clear relationship between these three estimated coefficients. These results continue to hold when data from each one of the locations are analyzed separately.

As univariate linear relationships have not been found, a multivariate linear analysis has been performed. Specifically, it has been analyzed whether there exists a group of independent (explanatory) variables which jointly explain the dependent variable whose data are estimations of one of the coefficients in the proposed ARMA model. The independent variables which have been used are: monthly mean values of the daily atmospheric clearness index, $K_{d,m}$, estimations of the other coefficients in the ARMA model and period of the year (the latter variable is described below). In all the series which have been used, the values of $K_{d,m}$ range between 0.32 and 0.74.

For each one of the three dependent variables

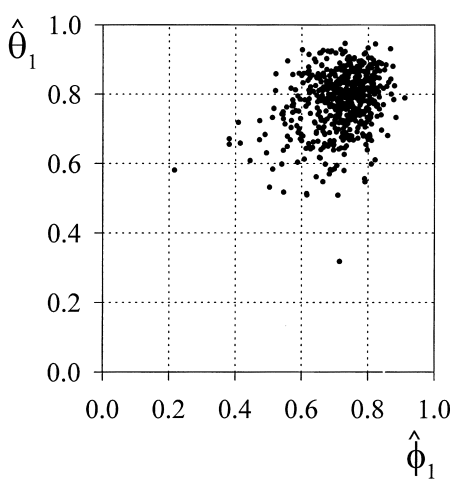

Fig. 1. The estimated autoregressive coefficient ($\hat{\phi}_1$) vs. the estimated moving-average coefficient ($\hat{\theta}_1$) of the multiplicative ARMA models for all monthly hourly series of global irradiation.

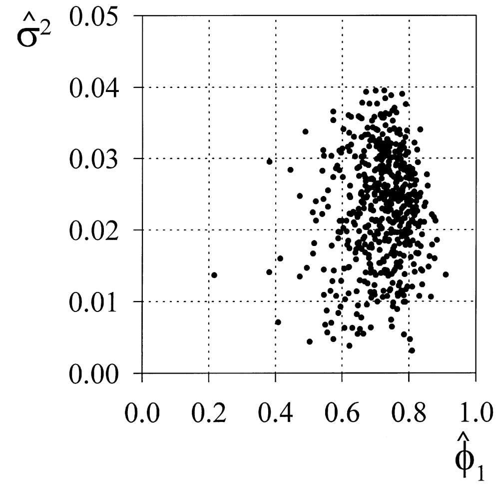

Fig. 2. The estimated autoregressive coefficient ($\hat{\phi}_1$) vs. the estimated variance of the white noise process ($\hat{\sigma}^2$) of the multiplicative ARMA models for all monthly hourly series of global irradiation.

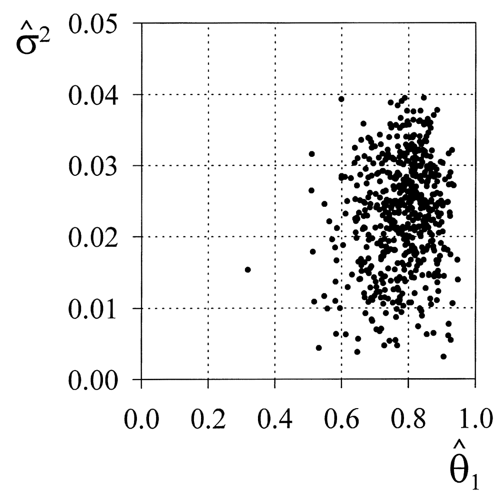

Fig. 3. The estimated moving average coefficient ($\hat{\theta}_1$) vs. the estimated variance of the white noise process ($\hat{\sigma}^2$) of the multiplicative ARMA models for all monthly hourly series of global irradiation.

(one variable for each estimated coefficient in the ARMA model), a regression analysis has been performed and it has been studied which independent variables are statistically significant. With this purpose, individual significance tests (as described, for example, in Hamilton, 1994, Chapter 8) have been used. These tests allows us to decide whether the independent variables in the model are useful to explain the variability of the dependent variable. In all cases, the significance level which has been used is 0.05.

The estimated values of the residual variance and autoregressive coefficient exhibit some annual dependence. It has been observed that they are systematically different, depending on the period of the year from which observations come. In order to include in the model this fact, two dummy variables have been used. These dummy variables allow us to group observations which correspond to the same period of the year.

The following model has been used to analyze the dependent variable "residual variance of the ARMA model":

$$\log(\hat{\sigma}^2)=\alpha_1 K_{\mathrm{d,m}}+\alpha_2 K_{\mathrm{d,m}}^2+\gamma_{\mathrm{m}}+\epsilon \quad (5)$$

where $\gamma_{\mathrm{m}}=\gamma_1$ if $5\leq$month≤ 8 and $\gamma_{\mathrm{m}}=\gamma_2$ otherwise, and ϵ is assumed to be a white noise process. The coefficients α_i $(i=1,2)$ and γ_j $(j=1,2)$ are estimated using multivariate regression methods (see, for example, Hamilton, 1994; the results obtained are described in Table 2. Using individual significance tests, all independent variables in the model turn out to be significant.

The model which has been used to analyze the dependent variable "estimated autoregressive coefficient" is:

$$\hat{\phi}_1=\alpha_1' K_{\mathrm{d,m}}+\alpha_2' K_{\mathrm{d,m}}^2+\alpha_3'\hat{\sigma}^2+\gamma_{\mathrm{m}}'+\epsilon \quad (6)$$

where $\gamma_{\mathrm{m}}'=\gamma_1'$ if $5\leq$month≤ 8 and $\gamma_{\mathrm{m}}'=\gamma_2'$ otherwise, and ϵ is assumed to be a white noise process. The variable "residual variance" has been included in the model because it proves to be significant. The coefficients in eqn (6) have been estimated using also regression techniques; the results obtained are shown in Table 3. Again, all independent variables in the model are statistically significant.

Table 2. Model fitting results for $\hat{\sigma}^2$ and coefficients for eqn (5)

$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
11.67	−13.52	−6.26	−6.06

Table 3. Model fitting results for $\hat{\phi}_1$ and coefficients for eqn (6)

$\hat{\alpha}_1'$	$\hat{\alpha}_2'$	$\hat{\alpha}_3'$	$\hat{\gamma}_1'$	$\hat{\gamma}_2'$
2.70	−3.26	−3.41	0.33	0.30

When studying the "estimated moving-average coefficient" as a dependent variable, however, none of the independent variables included in eqn (5) or eqn (6) proves to be statistically significant. Thus, there does not seem to be any relevant information in these variables which allows us to predict the behaviour of this dependent variable. For this reason, we have considered the values of this variable as independent and identically distributed observations from a univariate random variable, and have searched for a statistical distribution which adequately fits these observations. Among all distribution functions which are frequently used, the Weibull distribution is, in this case, the one which yields the best results for the estimated moving-average coefficients. This distribution function depends on two parameters (which will be denoted a and b) and its density function is:

$$f(x)=ab^{-a}x^{a-1}\exp[-(x/b)^a], \quad x>0. \quad (7)$$

When the estimated moving-average coefficients are used as observations, maximum likelihood techniques yield the following estimates of parameters a and b:

$$\hat{a}=11.25\ \hat{b}=0.816$$

Using goodness of fit tests (Kolmogorov-Smimov and χ^2, see, for example, Gibbons, 1971, chapter 4), it is accepted as the null hypothesis that the distribution from which these values come from is Weibull (significance level$=0.05$). In Fig. 4, an histogram of the variable "estimated moving-average coefficients" and the density function of the Weibull distribution with parameters â and $\hat{b}$ are depicted. If

$$F(x)=1-\exp[-(x/b)^a]$$

is the Weibull cumulative probability function, the variable:

$$U=F(x)$$

is uniformly distributed in [0,1]. Thus

$$W=b[\log(1/1-U)^{1/a}]$$

is a random number from a Weibull distribution

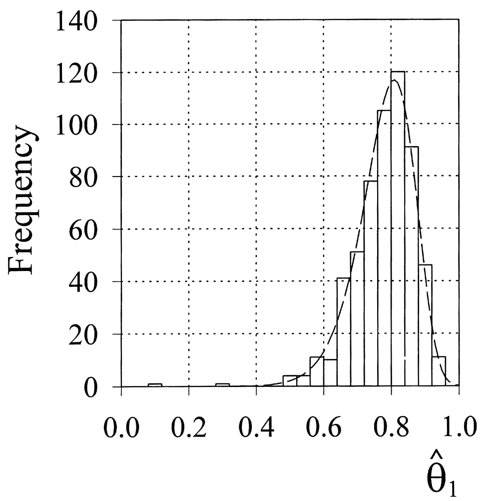

Fig. 4. A frequency histogram of estimated moving-average coefficients ($\hat{\theta}_1$) and the Weibull density function fitted for these values.

with density as given in eqn (7) (see Appendix A for a proof of this statement).

5. GENERATING NEW SERIES

All the results obtained in the previous section will be used to obtain simulated hourly series of global irradiation, $G'_{\mathrm{h}}(t)$. Monthly average values of daily global irradiation $G_{\mathrm{d,m}}$ must be available if the procedure which is described below is to be used. For each one of the locations these values can be taken directly from radiation atlas. Using these values, it is possible to compute, for each month, the value $K_{\mathrm{d,m}}$ With this value $K_{\mathrm{d,m}}$, it is possible to obtain for each month the value for $\hat{\sigma}^2$ using the eqn (5). The value for $\hat{\phi}_1$ can then be computed from the eqn (6). And the value for $\hat{\theta}_1$ is generated from a Weibull distribution with parameters $a=-11.25$ and $b=0.816$ as has been described at the end of the previous section.

With all these values, it is possible to generate a series of values $Y(t)$ which follows the proposed seasonal multiplicative ARMA model. With this purpose, the following steps must be performed:

(1) Generate observations $\epsilon(t)$ from a Gaussian white noise process whose variance is the value $\hat{\sigma}^2$ previously obtained
(2) Using an initial value, e.g. $Y(0)=0$, compute:

$$Y'(t)=\hat{\phi}_1 Y(t-1)+\epsilon(t) -\hat{\theta}_1\epsilon(t-s), \quad t=2,\ldots,Ns \quad (8)$$

where $\hat{\phi}_1,\hat{\theta}_1$ are values obtained as just described, N is the number of days in the month and s is the number of hours per day. Using the algorithm described in Appendix B, compute $X_{\mathrm{h}}(t)$.
Using these values and the expression of the maximum hourly global irradiation, obtain the following initial series of hourly irradiation, $[S'_{\mathrm{h}}(t)]$:

$$\begin{aligned} S'_h(t) &= G_{\mathrm{h,max}}(t)X'_{\mathrm{h}}(t) \\ &= 3.96[\cos\theta_{\mathrm{z}}(t)]^{1.05}X'_{\mathrm{h}}(t) \end{aligned} \quad (9)$$

Multiply the series obtained with eqn (9) by a coefficient in such a way that the resultant series has the same mean value as the recorded one; that is to say, obtain $G'_{\mathrm{h}}(t)$ using the expression:

$$G'_{\mathrm{h}}(t)=S'_{\mathrm{h}}(t)\frac{G_{\mathrm{h,m}}}{S'_{\mathrm{h,m}}} \quad (10)$$

where:

$$G_{\mathrm{h,m}}=\frac{G_{\mathrm{d,m}}}{s}$$

$$S'_{\mathrm{h,m}}=\frac{\sum_{t=1}^{N\cdot s} S'_{\mathrm{h}}(t)}{N\cdot s}$$

The value of the coefficient $S'_{\mathrm{h,m}}$ ranges from 0.85 to 1.12. It has been observed that for some values (approximately, 1%), the hourly irradiation value which is obtained is greater than the clear sky maximum. However, these values have been kept in the generated series, because in some recorded series some values which are greater than the clear sky maximum also appear—this happens because the latter is obtained using an empirical fit. The modification in Step 5 makes it possible to simplify the method to obtain the series $[X'_h(t)]$ which is proposed in Mora-López and Sidrach-de-Cardona (1997). It is worth noting that this modification is based on the use of a known parameter $G_{\mathrm{d,m}}$ and that the statistical properties of the series (autocorrelation function and autopartial correlation function) are not altered, i.e., the relationship between the values at a specific hour and at the following hour and the relationship between the values at a specific hour and at the same hour in the following day remain unchanged.

6. RESULTS

Following the steps described above, new hourly series of global irradiation have been generated for all locations under consideration.

For each simulated monthly series, plots with the resultant values have been depicted and in all cases they seem to be very similar to the plots obtained with the real series. For instance, in Figs. 5 and 6 the recorded and simulated series of hourly irradiation for April, 1997 in Oviedo are shown. In these figures it is possible to see that various patterns of day can be reproduced with the proposed model. Obviously, in the simulated and recorded series it is not possible to achieve that the sequence of days takes place in the same period; but it is possible to achieve the appearance of some similar patterns during the month. In Fig. 7 the

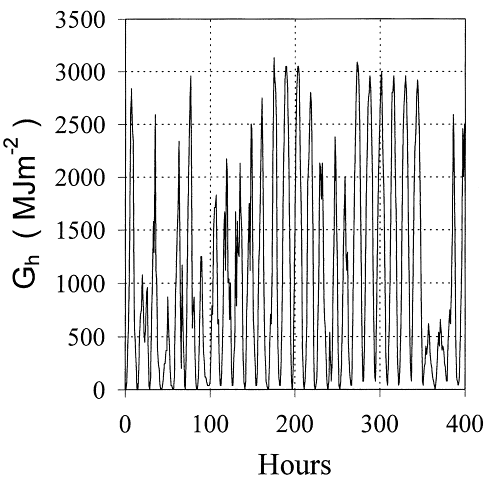

Fig. 5. The recorded hourly series of global irradiation. Data: Oviedo, April 1977.

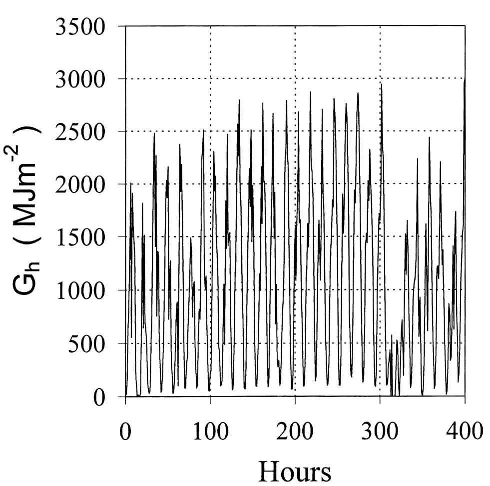

Fig. 6. The generated hourly series of global irradiation. Data: Oviedo, April 1977.

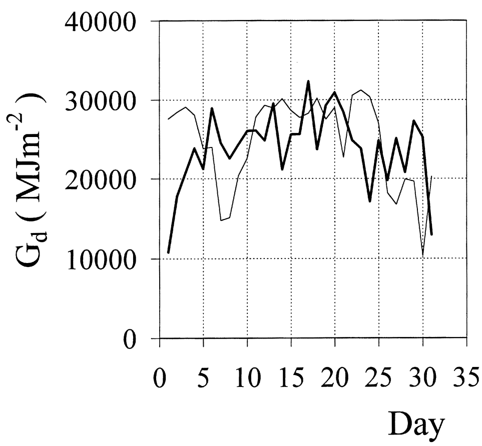

Fig. 7. The recorder (———) and simulated (———) values of daily global irradiation for May in Madrid.

recorded and simulated values for some days are shown. With these values it is possible to see, with greater detail, the evolution of both series, for May in Madrid. In Fig. 8 some examples of single days, for March in Oviedo, are shown; in this figure the hourly fluctuations could be appreciated.

Obviously, the mean values of recorded and simulated series coincide because this is a restriction which is imposed in the process of generation of the new series (see step 5 above). On the other hand, using statistical tests, the hypothesis that both series have the same variance is not rejected (significance level = 0.05). The sample autocorrelation functions and partial autocorrelation functions of the series obtained after removing the trends in the simulated series (using eqn (1) and eqn (3)) confirm that it is possible to identify these resultant simulated series with seasonal multiplicative

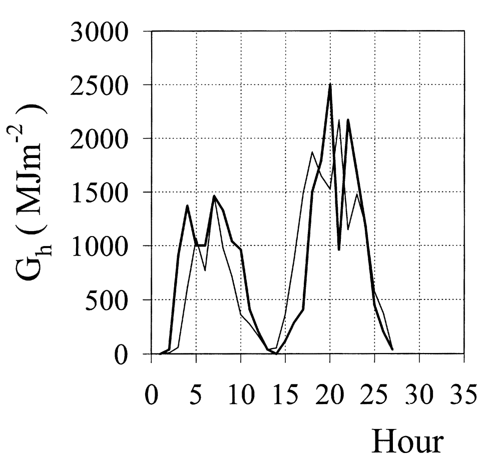

Fig. 8. The recorder (———) and simulated (———) series of hourly global irradiation. The examples are of single days in March in Oviedo.

ARMA(1,0)x(0,1)$_s$ models, as it could be expected, because this is precisely the statistical model which was used to generate the simulated series.

Finally, the frequency histograms of recorded and simulated series of two parameters have been also analyzed: the hourly clearness index and the hourly irradiation. The frequency histograms have been obtained for each month of the year, using all recorded and simulated series for that month throughout all years. That is to say, for each month in each location the following series have been compared:

$$\{K_h(t)\}_i \ \{K'_h(tc)\}_i \quad i=1,\ldots,Y$$
$$\{G_h(t)\}_i \ \{G'_h(t)\}_i \quad i=1,\ldots,Y \qquad (11)$$

In both cases, the histograms have been compared using the two-sample Kolmogorov-Smimov test-statistic (see, for example, Gibbons, 1971, chapter 7); the null hypothesis that the underlying model for both series is the same has never been rejected (significance level = 0.05). Cumulative probability distribution functions for the recorded and simulated series as in eqn (11) are shown: in Fig. 9 for the series of hourly clearness index (data from Madrid, May, 1979–1986); in Fig. 10 for the series of hourly irradiation (data from Madrid, January, 1979–1986)

We have also compared daily series of global irradiation obtained from recorded and simulated data of hourly exposure. In this case, using statistical tests, the following null hypotheses have never been rejected (significance level = 0.05): the mean values of these series are the same; the variances of these series are the same; and the distribution function of these series are the same (in the latter case, the two sample Kolmogorov-Smimov test statistic was used). In Fig. 11 the cumulative probability distribution functions for recorded and simulated series are shown (data from Madrid, 1979–86).

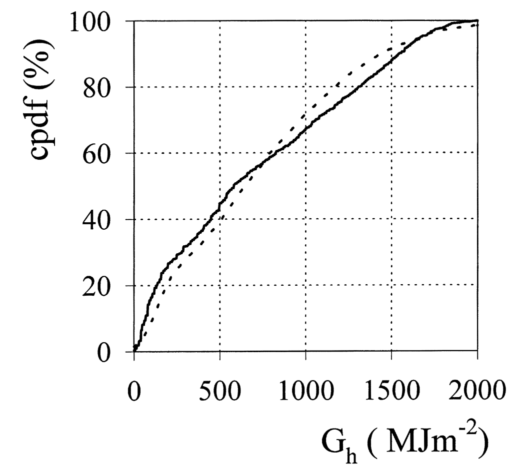

Fig. 10. The cumulative probability distribution function for the recorder (———) and generated (— — —) hourly series of global irradiation. Data: Madrid, January, 1979–86.

7. CONCLUSIONS

The use of multiplicative ARMA models to generate hourly series of global irradiation requires the knowledge of the parameters contained in these models for each monthly series which is to be generated. This is a consequence of the non-universality of the underlying param-

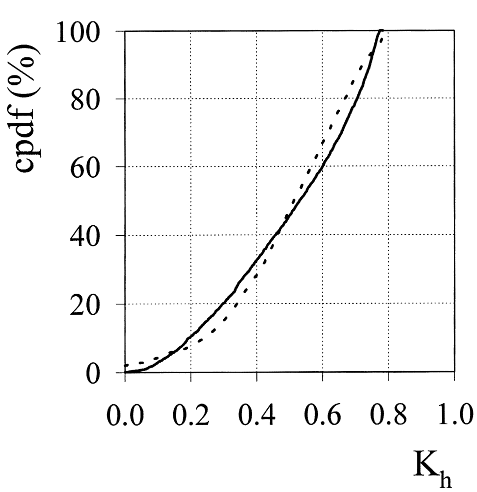

Fig. 9. The cumulative probability distribution function for the recorded (———) and simulated (— — —) series of hourly clearness index. Data: Madrid, May, 1979–86.

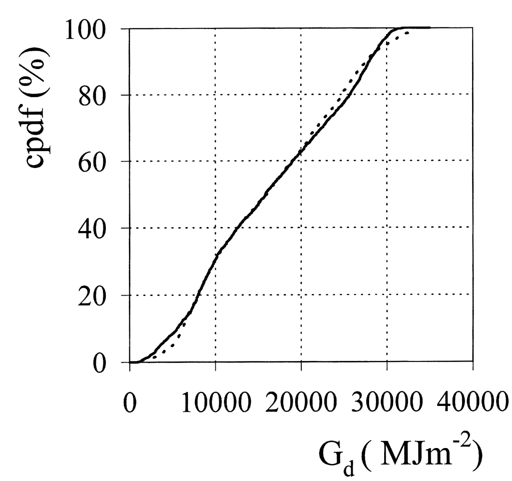

Fig. 11. The cumulative probability distribution function for the recorder (———) and generated (— — —) daily series of global irradiation. Data: Madrid, 1979–86.

eters of the models, i.e., these parameters vary from one location to another and from one month to another. In Mora-López and Sidrach-de-Cardona (1997), recorded series of hourly values of global irradiation were used to estimate these parameters; hence, this method cannot be applied directly when recorded series are not available. This is the reason why this work focuses on the analysis of the relationship between the parameters of the model and the available data in most locations. Expressions based on the results of this analysis can be used to obtain the various coefficients in the multiplicative ARMA models. The only required input value is the monthly mean value of daily global irradiation. In this paper a new method of obtaining hourly series is also proposed; this method simplifies the method proposed in the previous paper by Mora-López and Sidrach-de-Cardona (1997), which is specified in Appendix B. Thus, using only the hourly monthly mean value of global irradiation and following the steps described above in this paper, it is possible to obtain new hourly series of global irratiation.

In order to check the validity of the methods which are proposed to estimate the unknown parameters of the multiplicative ARMA models, synthetic hourly series of clearness index and of global irradiation have been obtained, and it has been verified that the generated series, for both parameters, have the same statistical characteristics as the real series. In fact, it can be accepted that the generated series have the same mean, variance and cumulative probability distribution function as the real series (using 0.05 as significance level). Obviously, if the daily and seasonal trends of hourly series of global irradiation are removed, the sample autocorrelation function and partial autocorrelation function of the new series (without trends) are similar to the theoretical autocorrelation function and partial autocorrelation function of the multiplicative seasonal ARMA $(1,0)\mathrm{x}(0,1)_s$. It has also been checked that the daily series which are obtained from the hourly series generated with ARMA models have the same statistical properties as the real ones. That is to say, the mean values, variances and cumulative probability distribution functions of both the daily series obtained from recorded values and the daily series obtained from generated values are statistically equal.

The analysis required still remains however, to determine if the proposed model and the expressions which are proposed to estimate the parameters of this model continue to hold in locations where $K_{d,m}$ (monthly mean value of the daily atmospheric transparency index) is very different from the values which have been used in this work.

NOMENCLATURE

$\alpha_i,\alpha'_i,\gamma'_i,\gamma'_i$	coefficients estimated using regression techniques
ϕ_1	autoregressive coefficient of the ARMA(1,0) $\times(0,1)_s$ model
$\hat{\phi}_1$	estimated autoregressive coefficient
θ_1	moving average coefficient of the ARMA(1,0) $\mathrm{x}(0,1)_s$ model
$\hat{\theta}_1$	estimated moving average coefficient
θ_z	zenith angle
σ^2	variance of the white noise process $\epsilon(t)$
$\hat{\sigma}^2$	estimated variance
a,b	parameters of the density function of the Weibull distribution
$\epsilon(t)$	white noise process
$G_{d,m}$	monthly average of daily global irradiation
$G_h(t)$	hourly global irradiation at hour t
$G'_h(t)$	generated hourly global irradiation at hour t
$G_{h,max}(t)$	maximum hourly global irradiation
G_{dy}	annual average values of daily global irradiation
$K_{d,m}$	monthly average of daily clearness index
K_h	hourly clearness index
K'_h	generated hourly clearness index
N	number of days in the month
$S'_h\ (t)$	initial generated series of hourly global irradiation
$S'_{h,m}$	re-scaling factor, for the initial generated series of hourly irradiation
s	number of hours, per day
$X_h(t)$	hourly global irradiation divided by maximum hourly irradiation
$X'_h(t)$	generated hourly global irradiation divided by maximum hourly irradiation
$Y(t)$	differenced values of $X_h(t)$
$Y'(t)$	generated differenced values of $X'_h(t)$

REFERENCES

Aguiar R. J., Collares-Pereira M. and Conde J. P. (1988) Simple procedure for generating sequences of daily radiation values using a library of Markov Transition Matrix. *Solar Energy* **40,** 3, 269–279.

Aguiar R. and Collares-Pereira M. (1992a) Tag: A time-dependent autoregressive, gaussian model for generating synthetic hourly radiation. *Solar Energy* **49,** 3, 167–174.

Aguiar R. and Collares-Pereira M. (1992b) Statistical properties of hourly global radiatior. *Solar Energy* **48,** 3, 157–167.

Bendt P., Collares-Pereira M. and Rabl A. (1981) The frequency distribution of daily insolation values. *Solar Energy* **27,** 1–5.

Boch G., Boileau E. and Bernard C. (1981) Modelisation of the hourly solar global irradiation in Trappes (France) by ARMA processes. LSS-EEE, 2418–2423.

Box J. E. P and Jenkins G. M. (1976) *Time Series Analysis, Forecasting and Control*. Holden-Day, p. 575.

Brinkworth B. J. (1977) Autocorrelation and stochastic modelling of insolation series. *Solar Energy* **19,** 343–347.

Gibbons J. D. (1971) *Nonparametrics Statistical Inference*. McGraw-Hill, Tokyo.

Graham V. A. and Hollands K. G. T. (1990) A method to

generate synthetic hourly solar radiation globally. *Solar Energy* **44,** 6, 333–341.
Guinea D., Mora Ll. and Palomo E. (1988) Análisis estadístico de series de exposición horaria de radiación solar, 4° Congreso Ibérico, 2° Iberoamericano de Energiá Solar, Oporto, Portugal.
Hamilton, J. D. (1994) *Time Series Analysis*. Princeton, New Jersey.
Mora-López Ll. (1995) Caracterización y generación de secuencias horarias de radiación global. Tesis Doctoral. Universidad Complutense de Madrid Abril.
Mora-López Ll. and Sidrach-de-Cardona M. (1997) Characterization and simulation of hourly exposure series of global radiation. *Solar Energy* **60,** 5, 257–270.
Palomo E. (1989) Hourly solar radiation time series as first-order Markov chains, Proc.ISES Solar World Congress, Kobe, Japan.

APPENDIX A

Generating random numbers with Weibull distribution from random numbers with Uniform distribution.

In this appendix we describe how to generate random numbers from a Weibull distribution following well known statistical procedures.

If U is a random variable with uniform distribution in (0,1) and W is defined as:

$$W=b[-\log(1-U)]^{-1/a}$$

then W is always in $(0,+\infty)$ and if x is a real number in $(0,+\infty)$ then the distribution function of W in x is:

$$F(x)=\Pr(W\leq x)=\Pr\{b[-\log(1-U)]^{-1/a}\leq x\}$$
$$=\Pr\{U\leq 1-\exp[-(x/b)^a]\}=1-\exp[-(x/b)^a]$$

The density function of W is the derivative of this expression with respect to x; but the derivative which is obtained from this expression is precisely the function $f(x)$ which appears in eqn (7).

APPENDIX B

Algorithm to calculate the series $[X'_{\mathrm{h}}(t)]$

First, observe that, as the variable t is a function of the hour and day, the series $[Y'(t)]$ can be written as:

$$Y'(t)\equiv Y'(h,d)\ h=1,\ldots,s\ d=1,\ldots,N \quad \text{(B1)}$$

$$Y'(h,d)=\phi_1 Y'(h-1,d)-\theta_1\epsilon(h,d-1)+\epsilon(h,d) \quad \text{(B2)}$$

where s is the number of hours for the day and N is the number of days for each month.

Accordingly, the series $[X'\mathrm{h}(t)]$ can be written as:

$$X'(h,d)=Y'(h,d)+X'(h,d-1)\ h=1,\ldots,s\ d=1,\ldots,N \quad \text{(B3)}$$

But s initial values $X'(h,0)$ $(h=1,\ldots,s)$ are required if eqn (B3) is to be used. These s initial values must be chosen in such a way that all values in $[X'_h(t)]$ are greater than zero and less than one. In order to be sure that this condition is met, these s initial values are obtained with an iterative process which is now described.

First, a series $[Z_0(h,d)]$ is obtained as follows:

$$Z_0(h,0)=0\ h=1,\ldots,s \quad \text{(B4)}$$

$$Z_0(h,d)=Y'(h,d)+Z_0(h,d-1)\ h=1,\ldots,s\ d=1,\ldots,N \quad \text{(B5)}$$

Then, the minimum and maximum values which appear in this series for each hour are searched:

$$m(h)=\min[Z_0(h,j),j=0,\ldots,N] \quad \text{(B6)}$$

$$M(h)=\max[Z_0(h,j),j=0,\ldots,N] \quad \text{(B7)}$$

Then it is checked whether:

$$\mathrm{ABS}[m(h)]<1-M(h) \quad \text{(B8)}$$

If eqn (B8) is not satisfied at an hour, then the second minimum is set as $m(h)$ and eqn (B8) is checked; if it is not satisfied either, then the second maximum is set as $M(h)$ and eqn (B8) is checked again. The process continues until values $m(h)$ and $M(h)$ which satisfy eqn (B8) are found for each hour. Usually, the minimum and maximum values or second minimum, second maximum values satisfy eqn (B8).

In order to reproduce the trend observed in the series $[X_{\mathrm{h}}(t)]$, a new series $[X'_h(h,d)]$ is then obtained as follows:

$$X'_{\mathrm{h}}(h,0)=1-M(h) \quad \text{(B9a)}$$

$$X'_{\mathrm{h}}(h,0)=\mathrm{ABS}[m(h)]+\frac{[1-M(h)]-\mathrm{ABS}[m(h)]}{2} \quad \text{(B9b)}$$

$$X'_{\mathrm{h}}(h,0)=\mathrm{ABS}[m(h)] \quad \text{(B9c)}$$

$$X'_{\mathrm{h}}(h,d)=Y'(h,d)+X'_{\mathrm{h}}(h,d-1)\quad h=1,\ldots,s\ d=1,\ldots,N \quad \text{(B10)}$$

where the expression:

eqn (B9a) is used for the 4 hours around noon: from $(s/2-2)$ to $(s/2+2)$

eqn (B9b) is used for the hours: from $(s/2-4)$ to $(s/2-2)$ and from $(s/2+2)$ to $(s/2+4)$

eqn (B9c) is used for the remaining hours.

The series $[X'_h(h,d)]$ can be written as:

$$X'_{\mathrm{h}}(t)=X'_{\mathrm{h}}(h,d)\quad t=1,\ldots,sN$$
h = 1,…,s d = 1,…,N