Stochastic models for sunshine duration and solar irradiation

P.K. Jain [a,*], E.M. Lungu [b]

[a] *Department of Physics, University of Botswana, Private Bag 0022, Gaborone, Botswana*
[b] *Department of Mathematics, University of Botswana, Private Bag 0022, Gaborone, Botswana*

Received 9 August 2001 ; accepted 11 November 2001

Abstract

Harmonic analysis of sunshine duration and solar irradiation measured at Sebele, Botswana is carried out. The data used consists of the monthly averages and the Julian-days averages of sunshine duration and solar irradiation sequences. This study involves splitting the time series into deterministic and stochastic components, and determining the proportion of the variance explained by each component. The stochastic component is analyzed for persistence using the Box and Jenkins technique. It is found that the stochastic component for monthly averages solar radiation series is best described by the second-order autoregressive Markov process, while that for Julian-days averages series has no memory.

1. Introduction

Sunshine duration and solar irradiation for any particular location are primarily determined by the rotational and orbital motions of the earth. In the absence of atmospheric effects both sunshine duration and solar irradiation for any location would be deterministic. This is evident from the well-known expressions available in the literature [1] for the calculation of the extraterrestrial radiation, day length and other related parameters of solar irradiance corresponding to any location on earth. However, atmospheric effects and processes such as cloud cover, pollution, etc. are seasonal as well as stochastic in nature, and they attenuate the sunshine

* Corresponding author.
E-mail addresses: jainpk@mopipi.ub.bw (P.K. Jain); lunguem@mopipi.ub.bw (E.M. Lungu).

duration and solar irradiation. This results in the introduction of stochasticity to both sunshine duration and solar irradiation for a terrestrial location. Therefore, solar irradiation and its duration for any location on earth may be visualized as consisting of a deterministic component and a stochastic component. In this paper sunshine duration data for 25 years (1976–2000) and solar irradiation data for 17 years (1976–1992) for Sebele, Botswana (latitude 24°34′ S; longitude 25°57′ E; altitude 994 m) are analyzed by decomposing them into periodic and stochastic components as described by Jain and Lungu [2]. The periodic component is subjected to harmonic analysis, and the stochastic component is analyzed by using the Autoregressive Integrated Moving Average (ARIMA) model [3] following the approach used by Lungu and Sefe [4] for the analysis of monthly runoff sequences for rivers in Botswana. While ARIMA models have been applied to hydrological modeling, to our knowledge these models have not been used for the analysis of solar radiation data. A model is fitted to the stochastic component. The analysis sheds light on the measure of the persistence and dependence pattern of the stochastic component. The significance of each component in explaining the variance of the meteorological sequences is identified.

2. Characteristics of solar radiation for Sebele

Sebele, 10 km north-east of the capital city Gaborone, is the earliest site in Botswana where measurements of solar radiation data were started in September/October 1975 by the Department of Agricultural Research, Ministry of Agriculture, Botswana. Sebele, with an average of 314 sunny days and 3320 hours of sunshine per year [5], is a good representative of the solar conditions in the southern part of Botswana. To date Sebele remains the only site for which data for both daily sunshine duration and solar irradiation on a horizontal surface are available, although sunshine duration is also measured by the Department of Meteorological Services at 14 geographically well-distributed synoptic stations throughout the country. Sunshine duration is measured using the Campbell–Stokes recorder, and the daily global irradiation is measured using a Kipp and Zonnen pyranometer. The pyranometer was calibrated annually against a standard precision pyranometer following the IGY procedure [6] under clear skies. The daily sunshine duration data available are up-to-date, with only about 2% of data points missing. Solar irradiation data, on the other hand, are a bit more scanty due to failure of the measuring instruments, have about 5% of data points missing for the period 1975 to 1992, have been measured for just a few months during 1993 to 1994, and have not been measured beyond 1994. Sunshine duration and solar irradiation data are organized for analysis as follows:

1. "total series", ${}^{m}Y_{\alpha,t}$ and ${}^{J}Y_{\alpha,t}$ — these denote the discrete values of the average monthly and Julian-day measurements, respectively, for $t = 1$ to 12 months and $\alpha = 1$ to N years; and
2. "monthly averages series", ${}^{m}M_{t}$ and ${}^{J}M_{t}$ — these denote the periodic variables for monthly averages and Julian-day averages, respectively, of the measured values

over the N years of observation for $t = 1$ to 12 months. Thus, $M_t = (1/N)\sum_{\alpha=1}^{N} Y_{\alpha,t}$

All averages are taken by omitting the missing data points, which in effect amounts to replacing the missing data points with the corresponding averages. Since the sunshine duration and solar irradiation are dealt with separately to fit a separate model to each series (mono-variate modeling), it is not essential to consider sunshine duration and solar irradiation for identical durations, and it is also not important to match the solar irradiation and sunshine duration series by omitting corresponding data points from both the series as is done in solar radiation modeling by authors such as Jain and Jain [7], where Angstrom-type relations are used to correlate sunshine duration and solar irradiation.

The four "monthly averages" M_t series are given in Table 1 and shown graphically in Fig. 1. From the table and figure, the following features concerning solar radiation in Botswana are evident. Solar irradiation displays its usual sinusoidal pattern, reaching a minimum during the winter season (June/July) and a maximum during the summer months (December/January), whereas sunshine duration does not display similar behavior. Sunshine duration fluctuates around a mean value of about 9.0 h per day throughout the year. This can be explained from the cloud cover and rain patterns in Botswana. The rainy season extends from November to March, and it is generally cloudy/hazy or raining in the late afternoon through the night to mid morning hours. This reduces the sunshine duration during the summer months, thereby

Table 1
Monthly averages (M_t series) of measured daily sunshine duration and solar irradiation for Sebele, Botswana

Month	Sunshine duration (h per day)		Solar irradiation (MJ/m² per day)	
	Period: 1976–2000		Period: 1976–1992	
	${}^{m}M_t$ series	${}^{d}M_t$ series	${}^{m}M_t$ series	${}^{d}M_t$ series
January	8.89	8.31	23.58	24.39
February	8.76	8.00	23.21	23.20
March	8.18	7.70	20.45	19.56
April	8.83	9.13	18.24	18.75
May	9.16	9.35	16.68	17.24
June	9.02	9.25	15.14	15.76
July	9.31	9.15	15.71	15.71
August	9.72	10.15	18.17	18.68
September	9.18	9.24	20.16	19.59
October	8.89	8.45	22.37	20.99
November	9.02	10.07	24.15	25.81
December	9.23	8.20	25.13	23.94

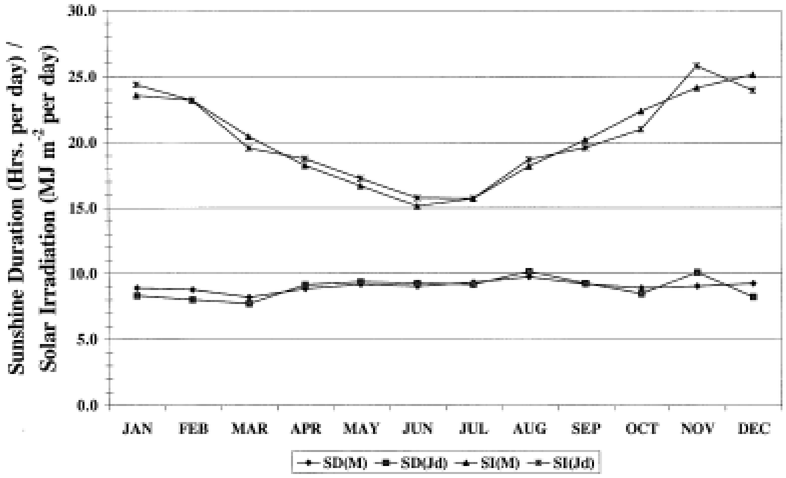

Fig. 1. Average monthly sunshine duration (SD) and solar irradiation (SI) (M_t series) for Sebele, Botswana (M, monthly averages series average; Jd, Julian-days averages series).

smoothing out the seasonal variation. However, the solar irradiation, which is most intense from mid morning until mid afternoon, does not lose its sinusoidal feature so much.

Table 2 gives the standard deviations, $\Delta Y_{\alpha,t}$ for the total series defined as:

$$\Delta Y_{\alpha,t} = \left[\frac{1}{12}\sum_{t=1}^{12}\frac{1}{N}\sum_{\alpha=1}^{N}\left(Y_{\alpha,t}-\frac{1}{N}\sum_{\alpha=1}^{N}Y_{\alpha,t}\right)^2\right]^{1/2} \quad (1)$$

and for the monthly averages series, ΔM_t, given as:

Table 2
Standard deviation for the eight data series of sunshine duration and solar irradiation

Series	Sunshine duration (h per day)		Solar irradiation (MJ/m² per day)	
	Monthly (^{m}) series	Julian-days (^{j}) series	Monthly (^{m}) series	Julian-days (^{j}) series
Total series, $Y_{\alpha,t}$	0.93	2.58	2.1	3.95
Monthly averages series, M_t	0.35	0.75	3.31	3.24

$$\Delta M_t = \left[\frac{1}{12}\sum_{t=1}^{12}\left(M_t-\frac{1}{12}\sum_{t=1}^{12}M_t\right)^2\right]^{1/2}. \tag{2}$$

With the exception of the solar irradiation ^JM_t, series, all other "Julian-days" series have considerably larger standard deviation compared with the corresponding "monthly" series. Thus the "Julian-days" series display a greater degree of fluctuation (stochasticity) as compared with the corresponding "monthly" series. This feature is also confirmed from the present analysis, and is discussed in greater detail later.

3. The ARIMA models

The general non-seasonal ARIMA model is autoregressive to order p and moving-average to order q, and operates on the dth difference of Z_i, where $\{Z_i\}$ are time series values for $i = 1, 2, ..., N$ and N is the number of observations. Defining

$$B^s Z_i = Z_{i-s}, \quad \nabla_s = (1-B^s), \quad \nabla_s^d = (1-B^s)^d, \tag{3}$$

where $d = 0, 1, 2, ..., B$ is the backward shift operator, s is the period of the season ($s = 12$ for the present case) and ∇ is the difference operator, the general non-seasonal ARIMA model can be written as:

$$\phi_p(B)\nabla^d Z_i = \theta_q(B)a_i, \tag{4}$$

where $\{a_i\}$ are residuals, and

$$\phi_p(B) = 1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p$$

and

$$\theta_q(B) = 1-\theta_1 B-\theta_2 B^2-...-\theta_q B^q$$

are polynomials of order p and q, respectively. Where an observation Z_i of a particular month has some relation with the observation made in the same month of the previous year, Eq. (4) can be modified [3] to account for the seasonal dependence. This yields

$$\phi_P(B^s)\nabla_s^D Z_i = \theta_Q(B^s)e_i, \tag{5}$$

where $\{e_i\}$ are normal random deviates, and

$$\phi_P(B^s) = 1-\phi_1 B^s-\phi_2 B^{2s}-...-\phi_P B^{Ps}$$

and

$$\theta_Q(B^s) = 1-\theta_1 B^s-\theta_2 B^{2s}-...-\theta_Q B^{Qs}$$

are the seasonal autoregressive and moving-average operators of order P and Q, respectively. As e_i is not necessarily independent of e_{i-j}, $j = 1, 2, ...$, we propose the following relation for the e values:

$$\phi_p(B)\nabla^d e_i = \theta_q(B)a_i. \tag{6}$$

Combining Eqs. (5) and (6) gives the general multiplicative seasonal ARIMA model of the order $(p, d, q) \times (P, D, Q)$ of the form:

$$\phi_P(B^s)\phi_p(B)\nabla_s^D\nabla^d Z_t = \theta_Q(B^s)\theta_q(B)a_t. \tag{7}$$

4. Objectives

The objectives of this paper are: (1) to identify the order (p, d, q) for a non-seasonal model, or the order $(p, d, q) \times (P, D, Q)$ for a seasonal model, that provides a parsimonious representation for both the total series under consideration and the stochastic component e_i; (2) to determine the proportions of the variance of the total series for the sunshine duration and the solar radiation data explained by the periodic and stochastic component; and (3) to determine the persistence pattern (if any) of the stochastic component.

5. Analysis and results

5.1. Harmonic analysis and the periodic components

Let $Y_{\alpha,t}$ denote the discrete values of an observed meteorological time series (either sunshine duration or solar irradiation of either of the two sequences, namely the monthly and the Julian-days series) with t and α representing the tth interval of the αth year, with $t = 1, 2, ..., 12$ and $\alpha = 1, 2, ..., N$. In view of the fact that the solar phenomena are cyclic and stochastic in nature, we can decompose it into a periodic component, M_t, and a stochastic component, e_i, as:

$$Y_{\alpha,t} = M_t + e_i, \tag{8}$$

where $i = 1, 2, ..., m$ and m ($= 12N$) is the number of observations. The periodic, deterministic component is reflected in the monthly mean values, $t = 1$ for January and $t = 12$ for December, and the stochastic component e_i is with mean zero and variance σ_e^2. The periodic component may be expressed as:

$$M_t = A_0 + \sum_{j=1}^{6} [A_j \cos(\pi jt/6) + B_j \sin(\pi jt/6)], \tag{9}$$

where the harmonic coefficients A_j and B_j are given by:

$$A_0 = \frac{1}{12}\sum_{t=1}^{12} M_t, \quad A_j = \frac{1}{6}\sum_{t=1}^{12} M_t \cos\left(\frac{\pi jt}{6}\right), \quad B_j = \frac{1}{6}\sum_{t=1}^{12} M_t \sin\left(\frac{\pi jt}{6}\right). \tag{10}$$

Table 3 gives the significance and variance of various harmonics of ${}^mY_{\alpha,t}$, ${}^mM_t$, ${}^jY_{\alpha,t}$ and ${}^jM_t$ series of sunshine duration. Similar data for solar irradiation are displayed in Table 4.

Table 3
Significance and variance explained by various harmonics of the sunshine duration series

Harmonics	Significance (%)				Variance (%)			
	Monthly series		Julian-days series		Monthly series		Julian-days series	
	${}^{m}Y_{\alpha,t}$ series	${}^{m}M_t$ series	${}^{J}Y_{\alpha,t}$ series	${}^{J}M_t$ series	${}^{m}Y_{\alpha,t}$ series	${}^{m}M_t$ series	${}^{J}Y_{\alpha,t}$ series	${}^{J}M_t$ series
First	14.3	49.9	–	47.7	14.4	47.0	–	44.0
Second	–	15.2	–	6.5	–	14.3	–	5.9
Third	–	18.8	–	7.7	–	17.7	–	7.1
Fourth	–	9.5	–	20.8	–	9.0	–	19.1
Fifth	–	1.3	–	16.3	–	1.2	–	15.1
Sixth	–	5.3	–	1.0	–	4.9	-	0.9
Other	–	–	5.6 (31st)	–	–	–	5.1 (31st)	–
Stochastic	85.7	–	94.4	–	85.6	5.9	94.9	7.9
Total	100	100	100	100	100	100	100	100

Table 4
Significance and variance explained by various harmonics of the solar irradiation series

Harmonics	Significance (%)				Variance (%)			
	Monthly series		Julian-days series		Monthly series		Julian-days series	
	$^{m}Y_{\alpha,t}$ series	$^{m}M_{t}$ series	$^{J}Y_{\alpha,t}$ series	$^{J}M_{t}$ series	$^{m}Y_{\alpha,t}$ series	$^{m}M_{t}$ series	$^{J}Y_{\alpha,t}$ series	$^{J}M_{t}$ series
First	–	98.6	–	93.1	–	90.5	–	85.4
Second	–	0.3	–	0.3	–	0.3	–	0.5
Third	–	0.1	–	0.3	–	0.2	–	0.2
Fourth	–	0.5	–	4.2	–	0.5	–	3.9
Fifth	–	0.2	–	2.0	–	0.2	–	1.8
Sixth	–	0.3	–	0.1	–	0.3	–	0.2
Other	82.3 (17th)	–	53.4 (17th)	–	39.8 (17th)	–	74.3 (17th)	–
Stochastic	17.7	–	46.6	–	60.2	–	25.7	–
Total	100	100	100	100	100	100	100	100

Subjecting the monthly averages and total series to harmonic analysis reveals that: (1) for the monthly averages series of sunshine duration, the first harmonic is the most significant as it explains about 45% of the variance of both the M_t series; (2) for the total sunshine duration series, the e_t component explains about 86% of the variance of the ${}^{m}Y_{\alpha,t}$ series and about 95% of the variance of the ${}^{J}Y_{\alpha,t}$ series; (3) for the monthly averages series of solar irradiation, only the first harmonic is significant as it explains about 91% of the variance of the ${}^{m}M_t$ series and about 85% of the ${}^{J}M_t$ series; and (4) for the total solar irradiation series the 17th harmonic is the most significant, as it explains about 40% of the variance of the ${}^{m}Y_{\alpha,t}$ series and 74% of the variance of the ${}^{J}Y_{\alpha,t}$ series. From this, it is evident that the e_t component explains about 60% of the variance of the ${}^{m}Y_{\alpha,t}$ series, and only 26% of the variance of the ${}^{J}Y_{\alpha,t}$ series of solar irradiation.

5.2. Stochastic components

The e_t sequences may now be calculated from Eqs. (8)–(10). Subjecting the e_t sequences to autocorrelation analysis shows the following. (1) For both the $Y_{\alpha,t}$ series of sunshine duration, stationarity is achieved for $d = 0$ and $D = 0$. For these parameters the autocorrelation functions mimic the autoregressive behavior. This is further confirmed by the partial autocorrelations corresponding to $d = 0$ and $D = 0$ (Fig. 2). (2) For both the $Y_{\alpha,t}$ series of solar irradiation, stationarity is achieved for $d = 2$ and $D = 0$. The autocorrelation functions mimic the autoregressive behavior. This is confirmed by the partial autocorrelation functions, which display

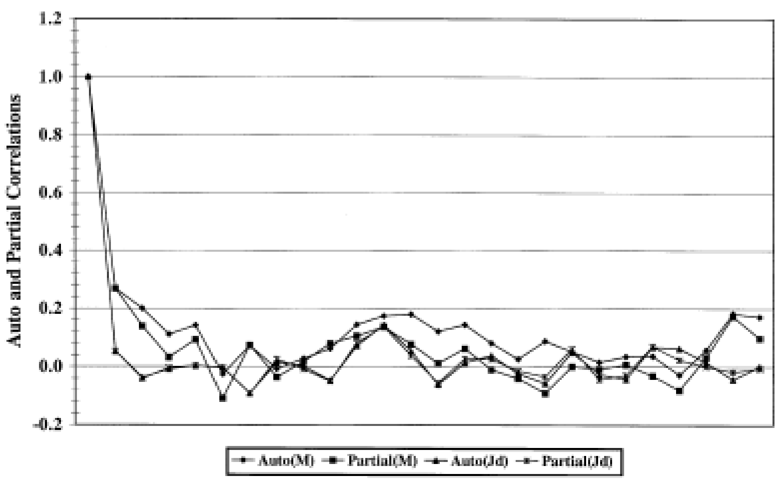

Fig. 2. Auto and partial correlations for the e_t sequences for total sunshine duration series ${}^{m}Y_{\alpha,t}$ and ${}^{J}Y_{\alpha,t}$ (M, monthly series; Jd, Julian-days series).

Table 5
Parameters and model statistics for total sunshine duration and solar irradiation series

Data series		Autoregressive parameters		Model	Portmanteau statistics	Critical value of statistics	Remark: model
		ϕ_{21}	ϕ_{22}				
Sunshine duration	${}^{m}Y_{\alpha,t}$	0.35	0.2	AR(2)	21.6	35.0	Accept
	${}^{r}Y_{\alpha,t}$	–	–	White noise	8.2	35.0	Accept
Solar irradiation	${}^{m}Y_{\alpha,t}$	(−)1.30	(−)0.17	AR(2)	15.7	35.0	Accept
	${}^{r}Y_{\alpha,t}$	–	–	White noise	1.8	35.0	Accept

the behavior of a moving-average process. (3) The partial and autocorrelations for solar irradiation display a similar pattern as for sunshine duration shown in Fig. 2. (4) Compared with 95% confidence limits, few of the partial autocorrelations are significant for both sunshine duration (Fig. 2) and solar irradiation. (5) The analysis suggests that e_t sequences follow the autoregressive process of order 2 for the ${}^{m}Y_{\alpha,t}$ series of both sunshine duration and solar irradiation. (6) The corresponding e_t sequences of the ${}^{J}Y_{\alpha,t}$ series for both sunshine duration (Fig. 2) and solar irradiation display the characteristics of pure white noise. Thus, for the ${}^{m}Y_{\alpha,t}$ series, we obtain:

$$\begin{aligned} e_t &= (\phi_{21}B + \phi_{22}B^2)e_t + a_t \quad \text{(for sunshine duration)} \\ &= (\phi_{21}B + \phi_{22}B^2)\nabla^2 e_t + a_t \quad \text{(for solar irradiation)} \end{aligned} \tag{11}$$

whereas $e_t = a_t$ for the ${}^{J}Y_{\alpha,t}$ series of both sunshine duration and solar irradiation. The values of $\phi_{\cdot,j}$ may be determined from [3]:

$$\phi_{k+1,j} = \phi_{k,j} - \phi_{k+1,k+1}\phi_{k,k-j+1} \ (j = 1, 2, ..., k), \tag{12}$$

$$\phi_{k+1,k+1} = \frac{r_{k+1} - \sum_{j=1}^{k} \phi_{k,j} r_{k+1-j}}{1 - \sum_{j=1}^{k} \phi_{k,j} r_j},$$

where r_j, $j = 1, 2, ..., k$, are the autocorrelations. The values of $\phi_{\cdot,j}$ are given in Table 5. The autoregressive parameters ϕ_{21} and ϕ_{22} for the monthly sunshine duration averages indicate fairly uniform variability and dependence of the stochastic component. However, for monthly solar irradiation averages there is higher dependence on the first stochastic component (ϕ_{21}). This means that the month-to-month effects are more important than dependence on past events (i.e. two months ago). For Julian-day averages the e_t sequences are found to be stochastic with no memory at all on previous events.

Subjecting the a_t sequences for both monthly sunshine duration and solar irradiation series to autocorrelation analysis shows: (1) that the sequences are normally distributed; (2) compared with 95% confidence limits all autocorrelations of the a_t sequences are insignificant, as shown in Fig. 3 for the a_t sequences corresponding to sunshine duration; and (3) the AR(2) model for sunshine duration and solar irradiation [Eq. (11)] pass the portmanteau statistics test with the portmanteau statistics computed using the first 25 values of the autocorrelations. Hence, the a_t sequences can be considered to be indistinguishable from a white noise sequence. On this basis, it is concluded that Eq. (11) can be considered to be a valid representation for stochastic components e_t for both the total series analyzed.

6. Conclusions and discussion

It has been shown that the stochastic components play a major role as the variance of both the ${}^{m}Y_{\alpha,t}$ and the ${}^{J}Y_{\alpha,t}$ series for sunshine duration, and the ${}^{m}Y_{\alpha,t}$ series for

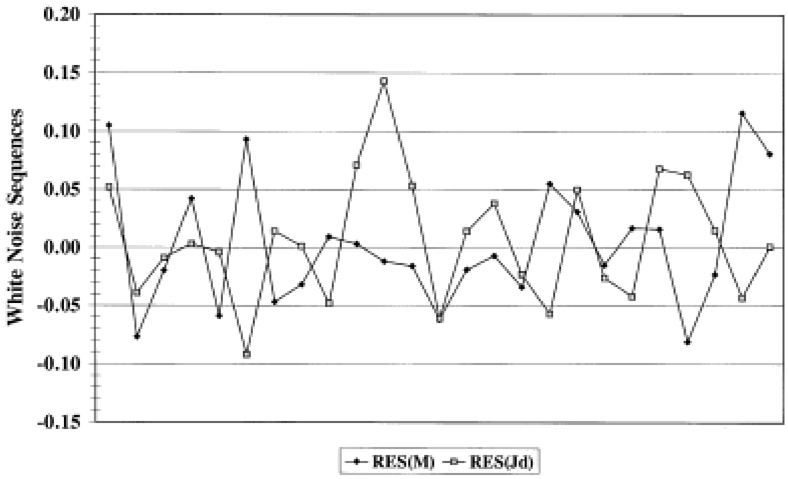

Fig. 3. Autocorrelation functions of the white noise (residues) for the total sunshine duration series ${}^{m}Y_{\alpha,t}$ and ${}^{J}Y_{\alpha,t}$ (M, monthly series; Jd, Julian-days series).

solar irradiation, are explained by this component. The values of the autoregressive parameters, ϕ_{2j}, for both sunshine duration and solar irradiation give a measure of persistence of the stochastic process. It indicates, for the monthly sunshine duration (${}^{m}Y_{\alpha,t}$ series), fairly uniform dependence on the previous two months and for monthly solar irradiation (${}^{m}Y_{\alpha,t}$ series) strong dependence only on the previous month. Both the Julian-days series (${}^{J}Y_{\alpha,t}$) of sunshine duration and solar irradiation indicate that the sequence ε_t is purely random without any memory. This result is not surprising since even during a month when there are plenty of sunshine days, the Julian day could be cloudy, partly cloudy or clear. This variability in Julian days is not smoothed for the month even if one took the average of Julian-days data over many years of observations (in the present study, 25 years in the case of sunshine duration and 17 years for solar irradiation), whereas the monthly averages for both sunshine duration and solar irradiation are fairly good representatives of the monthly average daily measure of these parameters from year to year. Thus Julian day is useful only where one needs to represent the monthly average quantities such as extraterrestrial solar irradiation, day length, etc. that are not influenced by atmospheric effects. This observation is supported by the fact (Table 4) that, for solar irradiation, the variance of the ${}^{J}Y_{\alpha,t}$ series is deterministic as it is explained mainly by the 17th harmonic whereas the stochastic component is pure white noise. Thus we conclude that Julian-days averages may not be used to represent the measured monthly averages of terrestrial sunshine duration and solar irradiation even if such averaging is carried over a reasonably long period. Furthermore, the conclusion that the stochastic components of the monthly series (${}^{m}Y_{\alpha,t}$) for both sunshine duration and solar irradiation have a memory of two months is in agreement with the traditional knowledge for this region, where the solar conditions during summer and winter months tend to be uniform over consecutive months. This can be used to simulate solar conditions with a lead time of two months. Such information is vital in climatic conditions such as Botswana's, where droughts are frequent and crop failures are common. Furthermore, the persistence pattern of solar radiation can be exploited to optimize the sizing of solar thermal and photovoltaic systems so that one may draw maximum benefits from the devices without having to oversize. Thus, we believe that this study sheds light on the characteristics of solar radiation and may help planners in agriculture and related industries in their planning.

Acknowledgements

Cooperation of the Department of Agricultural Research, Sebele, for providing the daily sunshine duration and solar radiation data used in the present analysis, is gratefully acknowledged. Thanks are due to Dr J. Prakash, Department of Basic Sciences, Botswana College of Agriculture, Sebele for tabulating the data.

References

[1] Duffie J, Beckman WA. Solar engineering of thermal processes. 2nd ed. New York: John Wiley and Sons, Inc, 1991.
[2] Jain PK, Lungu EM. Harmonic analysis of solar radiation data for Sebele, Botswana. In: Sayigh AAM, editor. Renewable energy: renewables — the energy for the 21st century (Proceedings of WREC-2000: World Renewable Energy Congress VI, Brighton, UK). UK: Pergamon; 2000:2575–8.
[3] Box GEP, Jenkins GM. Time series analysis: forecasting and control. 1st ed. San Francisco (CA): Holden-Day, 1970.
[4] Lungu EM, Sefe FTK. Stochastic analysis of monthly stream flows. J Hydrol 1991;126:171–82.
[5] Jain PK, Nijegorodov N, Kartha CG. Role of solar energy in development in Botswana. Int J Renewable Energy 1994;4:529–43.
[6] Radiation instruments and measurements. In: IGY instruction manual, vol. V, part VI. London: Pergamon Press, 1958.
[7] Jain S, Jain PC. A comparison of the Angstrom-type correlations and the estimation of monthly average daily global irradiation. Solar Energy 1988;40:93–8.