Journal of Agrometeorology 10(1) : 46 - 52 (June 2008)

Forecasting monthly wind speed for Udaipur region

S. R. BHAKAR and R. V. SINGH
Department of Soil and Water Engineering, College of Technology and Engineering, MPUAT, Udaipur-313001, Rajasthan. Email: srbhakar@rediffmail.com

ABSTRACT

Stochastic modelling for mean monthly wind speed of Udaipur (Rajasthan) was done using 26 years (1978–2003) data. The performed statistical tests indicated that the series of the monthly wind speed data is trend free. The periodic component can be represented by third harmonic expression. The stochastic components of the mean monthly wind speed follow fourth order Markov model. The correlation coefficient between generated and measured mean monthly wind speed series was 0.9995 and found to be highly significant 1 per cent level. The standard error (5.57 mm) is quite low. The regression equation is very near to 1:1 line. Therefore, developed model can be used for future prediction of monthly wind speed at Udaipur.

Keywords: Stochastic, auto correlation function, auto regression, wind speed.

Wind speed is an important weather parameter for estimation of crop water requirements. Frequently, it is required to estimate wind speed of places where measured wind speed data are not available. This type of series could be estimated by the sum of periodic series and stochastic series. Periodic component takes into account the portion, which repeats after certain duration. The stochastic component constituted by various random effects, cannot be estimated exactly.

MATERIALS AND METHODS

Study area and data

The study was conducted at the Department of Soil and Water Engineering, College of Technology and Engineering, MPUAT, Udaipur. The area comes under the sub-humid region of the agro-climatic zone IV-A of the state of Rajasthan, and is situated at 24°35' N latitude, 73°42'E longitude and at an altitude of 582.17 m amsl. The climate of an area is characterized by variation in the climatological parameters over the years and from one year to the next.

The wind speed and other data for a period of 26 years (1978-2003) were used in the study. Data were from Meteorological Observatory of the College of Technology and Engineering, Udaipur. (Bhakar, 2000).

The principal aim of the analysis is to obtain a reasonable model for estimating the generation process and its parameters by decomposing the original data series into its various components. Generally a time series can be decomposed into a deterministic component, which could be formulated in manner that allowed exact prediction of its value, and a stochastic component, which is always present in the data and cannot strictly be accounted for as it is composed of random effects. The time series, X(t), was represented by a decomposition model of the additive type for trend, periodic and stochastic component, as follows:

$$X_{(t)} = T_{(t)} + P_{(t)} + S_{(t)} \qquad ....(1)$$

To obtain the representative stochastic model of time series, identification and detection of each component of Equation (1) was necessary. A systematic identification and reduction of each component of $X_{(t)}$ was done, procedures of which are described below:

Trend component

For detecting the trend, a hypothesis of no trend was made. Turning point test and Kendall's rank correlation tests as suggested by Kottegoda (1980), were performed: If the calculated value of z is within its table value, then, it can be concluded that the trend is not present in the series.

Periodic component

The periodic component concerns an oscillating movement which is repetitive over a fixed interval of time (Kottegoda, 1980). The existence of $P_{(t)}$ was identified by the correlogram. The oscillating shape of the correlogram verifies the presence of $P_{(t)}$, with the seasonal period P. The time series $X_{(t)}$ was expressed in Fourier form as follows:

$$X_{(t)}=A_0+\sum_{K=1}^{\infty}\left[A_K\cos\left(\frac{2\pi Kt}{p}\right)+B_K\sin\left(\frac{2\pi Kt}{p}\right)\right] \quad ...(2)$$

where, $$A_o=\frac{1}{N}\sum_{t=1}^{N}x_{(t)} \quad ...(3)$$

$$A_K=\frac{2}{N}\sum_{t=1}^{N}x_{(t)}\,\mathrm{Cos}\left(\frac{2\pi Kt}{p}\right) \quad ...(4)$$

and $$B_K=\frac{2}{N}\sum_{t=1}^{N}x_{(t)}\,\mathrm{Sin}\left(\frac{2\pi Kt}{p}\right) \quad ...(5)$$

These coefficients were obtained by a least square fit of the data to the K^{th} harmonic components, then a least squares approximation can be given by the finite series.

$$P_{(t)}=A_o+\sum_{K=1}^{M}\left[A_K\,\mathrm{Cos}\left(\frac{2\pi Kt}{p}\right)+B_K\,\mathrm{Sin}\left(\frac{2\pi Kt}{p}\right)\right] \quad ...(6)$$

where, M is the number of significant harmonics (maximum, P/2). For later use, it was more convenient to use the alternate form for $P_{(t)}$ given as under:

$$P_{(t)}=A_o+\sum_{t=1}^{M}D_K\,\mathrm{Cos}\left(\frac{2\pi Kt}{p}-\theta_K\right) \quad ...(7)$$

where, $$D_K=\sqrt{A_K^2+B_K^2} \quad ...(8)$$

and $$\theta_K=\mathrm{Arc\ tan}\left(\frac{A_K}{B_K}\right) \quad ...(9)$$

Stochastic component

A stochastic model of the form of Autoregressive model, AR, was used for the presentation of the time series. In this model, the current value of the process is expressed as a finite, linear aggregate of values of the process and a variate that is completely random. This model was applied to the $S_{(t)}$ which was treated as a random variable i.e. deterministic components were removed and the residual was stationary in nature. Mathematically, an autoregressive model of order p, AR(p) can be written as:

$$S_{(t)}=\sum_{K=1}^{P}\phi_{P,K}\,S_{(t-K)}+a_{(t)}$$

$$=\phi_{p,1}S_{(t-1)}+\phi_{p,2}S_{(t-2)}+.....+\phi_{p,p}S_{(t-P)}+a_{(t)} \quad ...(10)$$

The fitting procedures of the AR(p) model to the meteorological parameters series involved selection of order (p) of the model.

Diagnostic checking of the model

Diagnostic checking concerns the verification for the adequacy of the fitted model. The residuals were examined for any lack of randomness. If the residuals were not random or were autocorrelated, the model has to be modified until the residuals become uncorrelated. The residual, $a_{(t)}$, record defined by the difference, estimated by Equation (11) was used in analyzing the closeness of fit of the formulated model.

$$a_{(t)}=S_{(t)}-\sum_{K=1}^{p}\phi_{p,k}\,S_{(t-k)} \quad ...(11)$$

RESULTS AND DISCUSSION

For testing the statistical characteristics of monthly wind speed series 26 years were analyzed to get mean monthly wind speed series, and statistical characteristics were estimated. The values varied from 2.70 kmph in December to 8.83 kmph in June. Mean value of wind speed was found to be 5.0 kmph. There is large variability among the mean monthly values of wind speed of different years. This was further confirmed by the estimated deviations. The standard deviation in the monthly wind speed values ranged from 0.463 to 1.395 kmph during entire year. The estimated variance indicates that the coefficient of variation ranges from 0.214 to 1.947. This signifies the importance of variability of monthly wind speed series. Since the values of variance significantly different from zero, it confirms that wind speed is mutually dependent.

Serial correlation coefficient

The lag one serial correlation coefficient of observed series was found to be 0.715. The respective confidence limits were estimated as 0.117 to -0.124

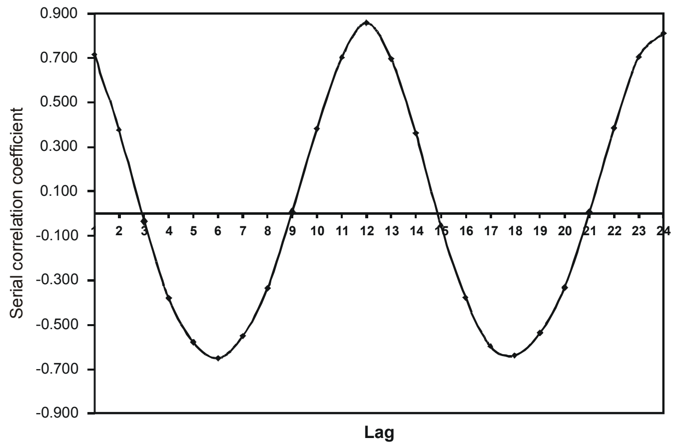

Fig. 1: Correlogram of annual wind speed Periodic Component

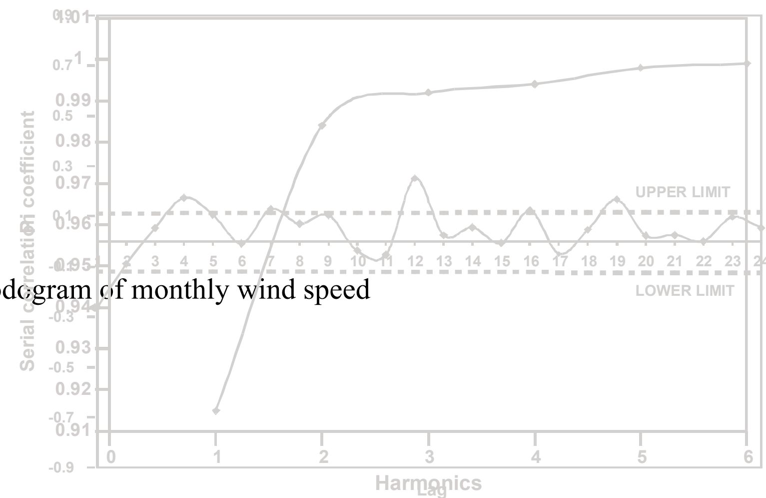

Fig. 2: Cumulative periodogram of monthly wind speed

Fig. 3: Correlogram up to lag 24 for residual series of mean monthly wind speed

Kottegoda (1980). The value of lag one serial correlation coefficient lies outside the range of confidence limits and is significantly different from zero. This again confirms that wind speed process is mutually dependent. From the analysis of coefficient of variation and serial correlation, it is confirmed that wind speed process is time variant and not an independent one. Thus the mean monthly wind speed time series may be modeled on stochastic theory. The mutual dependence of the observed wind speed series was also confirmed by the correlogram (Fig.1).

Trend component

For identification of trend components, annual wind speed series was used. The calculated values of test statistic (Z) are -0.34 and -3.87 for Turning point test and Kendall's correlation test respectively. The estimated values were within the 1 per cent level of significance. Hence the hypothesis of no trend was accepted.

Further from the turning point test total number of turning points in series were found to be 14. This indicates that the wind speed series is random. From the above analysis it is confirmed that the trend component in wind speed time series is absent and the observed series may be treated as trend free series.

To confirm the presence of periodic component in mean monthly wind speed series a correlogram was drawn (graphical representation of serial correlation coefficients (r_t) as function of lag. *l*) in which the value of r_t are plotted against respective value of *l* (Fig. 1). The resulting oscillating shape of the correlogram confirms the presence of periodic component in the monthly wind speed. Further, the correlogram has peaks at lags equal to 12 and at other multiples of it. The time span of periodicity was taken as 12 for use in harmonic analysis of periodic component.

For representing the periodic component of the wind speed series the numbers of significant harmonics were determined by analyzing by cumulative periodogram. Only first three harmonics are highly significant. Other harmonics are not significant and therefore ignored.

Parameters of periodic component

Using Equations (3), (4) and (5) the Fourier coefficients A_k and B_k were estimated. The amplitude, phase angle and explained variance for different harmonics were calculated by using Equations (8) and (9) respectively. These Fourier decompositions for monthly wind speed series reveals that the first three harmonics explain more than 91 per cent of variance.

Cumulative periodogram test

In this test a graphical procedure has been employed as criteria for obtaining the significant harmonics to be fitted in a periodic component. A cumulative periodogram was drawn between P_i and number of harmonics. The fast increase in Pi has been considered as a significant harmonics and the rest of harmonics were rejected.

Estimates of the mean square deviation and the cumulative periodogram Pi for mean monthly wind speed series was made and the plot of Pi against i is shown in Fig. 2. It can be observed that the first three harmonics appeared to be the periodic part of the fast increase and after that periodogram remains almost constant which may be treated as non-significant. The two criteria used to identify the number of significant harmonics to be used in modelling periodic component were found to be consistent, so first three harmonics are treated as significant contributing towards periodicity and remaining are considered as white noise.

For the first three harmonics the values of Fourier coefficients (A_1, A_2, A_3, B_1, B_1, B_3) were found to be -2.681, 0.729, -0.249 and 0.366, -0.139, -0.030 respectively. With these coefficients and using equations (6) the periodic component (P_t) resulting from periodic deterministic process may be mathematically expressed as:

$$P_t = 5.00 - 2.681\cos\left(\frac{2\pi t}{p}\right) + 0.366\sin\left(\frac{2\pi t}{p}\right) + 0.729\cos\left(\frac{4\pi t}{p}\right)$$

$$0.729\cos\left(\frac{4\pi t}{p}\right) - 0.139\sin\left(\frac{4\pi t}{p}\right)$$

$$-0.249\cos\left(\frac{6\pi t}{p}\right) - 0.030\sin\left(\frac{6\pi t}{p}\right) \quad \ldots(12)$$

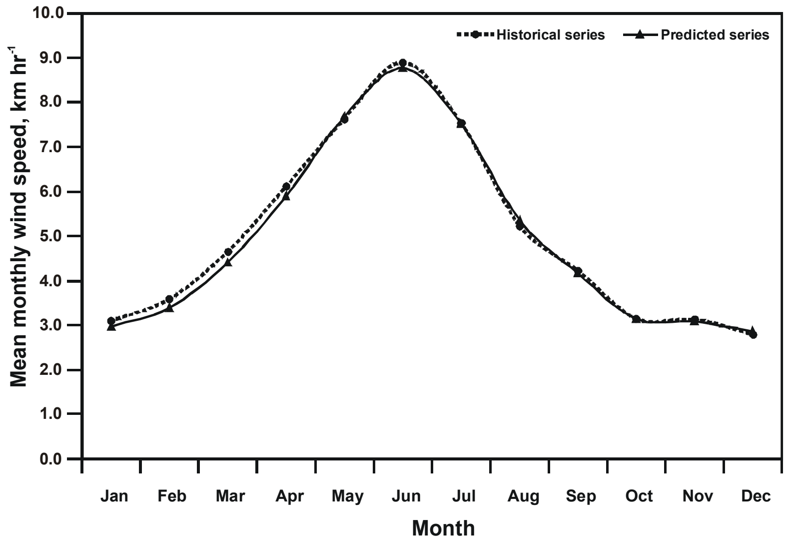

Fig. 4: Variation of generated mean monthly wind speed and measured mean monthly wind speed

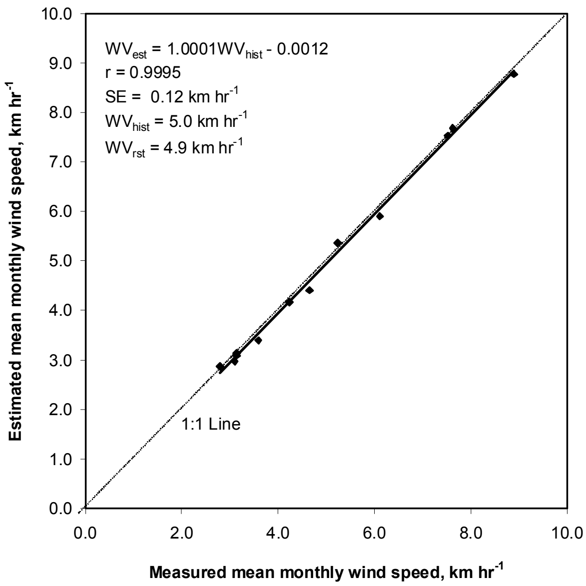

Fig. 5: Relationship between generated mean monthly wind speed ($Rain_{est}$) and measured mean monthly wind speed ($Rain_{hist}$)

The deterministic cycle component (Pt) was computed by using Equation (12) for all the values of t (t_{max}= 384). After estimating the periodic component it was subtracted from, historical time series. This process gave a new stationery series, S_t, resulting from stochastic non- deterministic process. The rest of the data were further analyzed to obtain non-deterministic stochastic component by fitting the autoregressive or Markov process of stochastic modeling.

Stochastic component

The presence of stochastic component was already confirmed by plotting the correlogram (Fig. 1) of observed series and analysis of serial correlation coefficient (SCC) and coefficient of variance (CV).

Table 1: Statistical parameters of the observed, generated and residual series of mean monthly wind speed

Parameters	Historical series	Generated series	Residual series
Mean, km hr^{-1}	5.0	5.0	0.0
SD, km hr^{-1}	2.237	2.123	0.302
Variance	5.004	4.508	0.091

Selection of model order

Residual variance method was used to determine the order of the model which may significantly represent the non-deterministic stationary stochastic component. Residual variance at different lags was computed with the help of equation (11). The minimum residual variance was obtained for order four after which they showed no definite trend. Using Equation (10) and the estimated autoregression coefficients the stochastic component of the monthly wind speed time series may be expressed as:

$S_t = 0.693\ S_{t-1} + 0.102\ S_{t-2} - 1.157\ S_{t-3} - 0.287\ S_{t-4} + a_t$ ..(13)

The non-deterministic stochastic component was estimated by using Equation (13) for all the values (t = 312).

The residual series of stochastic component

The residual series (a_t) which is random independent part of stochastic component was obtained with the help of equation (11), after removing the periodic and dependent stochastic parts from the historical series. The statistical analysis of the residual series confirms its normal distribution with mean which is almost equal to zero (mean 0.00 and SD 0.302 kmph). The values of statistical measures are presented is Table 1. The mean, SD of the historical and generated series are-almost same which shows closeness between historical and generated data.

Model structure

Since the observed monthly wind speed series was found to be a trend free series, the developed model describes the periodic-stochastic behaviour of the series and is a superimposition of harmonic deterministic process and third order autoregressive model. The mathematical structure of the additive model can now be represented as follows:

$$X_t = 5.00 - 2.681\cos\left(\frac{2\pi t}{p}\right) + 0.366\sin\left(\frac{2\pi t}{p}\right) + 0.729\cos\left(\frac{4\pi t}{p}\right) - 0.1$$

$$0.729\cos\left(\frac{4\pi t}{p}\right) - 0.139\sin\left(\frac{4\pi t}{p}\right) - 0.249\cos\left(\frac{6\pi t}{p}\right)$$

$$-0.030\sin\left(\frac{6\pi t}{p}\right) + 0.693\,S_{t-1} + 0.102\,S_{t-2}$$

$$-1.157\,S_{t-3} - 0.287\,S_{t-4} + a_t \qquad \ldots(14)$$

The first seven terms in the formulated model represented by Equation (14) constitute the deterministic part of the monthly wind speed time series. The eighth to eleventh terms represent the dependent stochastic component of the model where the current value of St depends on the weighted sum of observed preceding three values. The last term is the random independent part of the stochastic component. Using the developed model the average monthly wind speed series was generated for all the values (t = 312).

Diagnostic checking of wind speed model

The residuals obtained after fitting the formulated model was subjected to various analysis to test their adequacy for representing the time dependent structure of the mean monthly wind speed.

Sum of squares analysis

The sum of squares of residuals series were compared with sum of squares of deviations of observed values from their mean.. The value of coefficient of determination (R^2) was found to be 0.9991, which is nearly equal to unity. Thus, this leads to the conclusion that the developed model has a fair goodness of fit to generate the monthly wind speed series.

Serial correlation analysis

The serial correlation coefficients (SCC) for lags *l* (*l*= 1, 2, 3, ….24) were computed. The values of SCC against respective lags were then plotted to obtain a correlogram (Fig. 3) confidence limit at 1 per cent level. The correlogram is almost completely contained within these limits. Hence it may be treated to be non significant. This confirms that the residual series may be treated to be non significant. The residual series has a mean value of 0.00 (zero) and the variance of 1.464 which is approximately equal to variance 1/50th of historical series This leads to the conclusion that the residuals are independent and normally distributed. Further, it also confirms the randomness of the residuals.

Forecasting of mean monthly wind speed

Validation of generated 24 mean monthly wind speed series by developed stochastic model (Equation 14) was done by comparison with measured monthly wind speed series (Fig. 4), The correlation coefficient between generated and measured mean monthly wind speed series was found to be 0.9995 (Fig. 5). The correlation was tested by t test and found to be highly significant at 1 per cent level. The standard error (0.07 kmph) is quite low. The means of both the measured and generated monthly wind speeds were found to be 5.0 kmph. The regression equation is very near to 1:1 line. Therefore, Equation (14) can be used for future prediction of monthly wind speed.

REFERENCES

Bhakar, S. R. (2000). Modelling of evaporation and evapotranspiration under climatic conditions of Udaipur. Ph. D. thesis submitted to faculty of Agricultural Engineering, Department of Soil and Water Engineering, CTAE, MPUAT, Udaipur.

Kottegoda, N. T. (1980). Stochastic water resource technology. The MacMillan Press ltd., London. P: 384

Received : August 2005; Accepted : December 2007