Eti ENGINEERING AND TECHNOLOGY IN INDIA
Volume 3 | Issue (1) | April, 2012 | 1-6

RESEARCH ARTICLE

Stochastic modelling of maize evapotranspiration under the climatic conditions of Banswara

S.V. NIRMAL, S.R. BHAKAR, R.C. PUROHIT AND MAHESH KOTHARI

ABSTRACT : Stochastic modelling of maize evapotranspiration has done using 8 years (1998–2005) data. The performed statistical tests indicated that the series of the evapotranspiration data is trend free. The periodic component of evapotranspiration can be represented by second harmonic expression. The stochastic components of the evapotranspiration follow third order model. Validation of generated evapotranspiration series was done by comparison of generated evapotranspiration series and measured evapotranspiration series. The correlation coefficient between generated evapotranspiration series and measured evapotranspiration series was found to be 0.99. The correlation was tested by t-test and found to be highly significant at 1 per cent level. The standard error (0.13 mm) is quite low. The regression equation is very near to 1:1 line. Therefore, developed model can be used for future prediction of maize evapotranspiration series.

Key words : Stochastic, Auto correlation function, Auto regression

How to cite this Article : Nirmal, S.V., Bhakar, S.R., Purohit, R.C. and Kothari, Mahesh (2012). Stochastic modelling of maize evapotranspiration under the climatic conditions of Banswaras. *Engg. & Tech. in India,* **3**(1&2) : 1-6.

Article Chronicle : Received : *08.08.2011;* **Revised :** *05.11.2011;* **Accepted :** *15.01.2012*

INTRODUCTION

Evapotranspiration is an important parameter for estimation of crop water requirements. Frequently, it is required to estimate evapotranspiration of places where measured evapotranspiration is not available. The process of evapotranspiration is stochastic in nature. Usually, the deterministic models do not consider the random effects and may not represent the evapotranspiration quite accurately. On the other hand, the stochastic models are based on the time dependent variations and consider random effects involved in the process. Stochastic models explain the extent of dependence of a present observation on the past observations therefore; stochastic modelling of evapotranspiration may provide good insight and understanding of the processes for useful applications in water resources development. Keeping this in mind, the present study was undertaken with the objective of developing and validating the appropriate stochastic model for maize evapotranspiration.

MEMBERS OF RESEARCH FORUM

Address for correspondence :
S.V. NIRMAL, Department of Soil and Water Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, UDAIPUR (RAJASTHAN) INDIA
Email: sharad_vn@ayahoo.com

Coopted Authors :
S.R. BHAKAR, R.C. PUROHIT AND MAHESH KOTHARI, Department of Soil and Water Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, UDAIPUR (RAJASTHAN) INDIA

EXPERIMENTAL PROCEDURE

Location of the study area:

The study was conducted at the Agricultural Research Station, Borwat Farm, Banswara. The area comes under the sub-humid region of the agro-climatic zone IV-A of the state of Rajasthan, and is situated at $24^0 35'$ N latitude, $73^0 42'$E longitude and at an altitude of 582.17 m above mean sea level. The annual rainfall in this region is 646.6 mm and more than 80 per cent of this amount is received during the monsoon season alone, due to the influence of the southwest monsoon.

Collection of meteorological data:

The data of pan evaporation was collected from Meteorological Observatory of the College of Technology and Engineering, Banswara. Meteorological data for a period of 26 years (1978-2003) was used in the study.

Formulation of stochastic model:

The mathematical procedure adopted for formulation of a predictive model based on stochastic component has been discussed in the following sub-sections:

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 can not strictly be accounted for as it is made by random effects. The time series, $X_{(t)}$, was represented by a decomposition model of the additive type, as follows:

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

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:

The trend component describes the long smooth movement of the variable lasting over the span of observations, ignoring the short term fluctuations. The basic idea here was to study only $T_{(t)}$ while eliminating the effects of other components. It leads to use the total seasonal data, Z_t, for identification of $T_{(t)}$ so that other components were suppressed. 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, a plot of autocorrelation coefficients, r_1, versus lag *1*. The oscillating shape of the correlogram verifies the presence of $P_{(t)}$, with the seasonal period P, at the multiples of which peak of estimation can be made by a Fourier analysis followed by the tests for significant harmonics. The correlogram of the time series clearly shows the presence of the periodic variations indicating its detection. The time series $X_{(t)}$ was expressed in Fourier form as follows:

$$X_{(t)} = A_o + \sum_{K=1}^{\infty}\left[A_K \operatorname{Cos}\left(\frac{2\pi\pi K}{p}\right) + B_K \operatorname{Sin}\left(\frac{2\pi\pi K}{p}\right)\right] \quad (2)$$

where,

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

$$A_K = \frac{2}{N}\sum_{t=1}^{N} x_{(t)} \operatorname{Cos}\left(\frac{2\pi\pi K}{p}\right) \quad (4)$$

and

$$B_K = \frac{2}{N}\sum_{t=1}^{N} x_{(t)} \operatorname{Sin}\left(\frac{2\pi\pi K}{p}\right) \quad (5)$$

Stochastic component

The stochastic component is constituted by various random effects, which can not be estimated exactly. In case of maize evapotranspiration series various climatic parameters response to the value of component without changing the cyclicity itself and thus add randomness to the time series. 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:

$$S_{(t)} = \sum_{K=1}^{P} \varphi_{P,K}\, S_{(t-K)} + a_{(t)} \quad (6)$$

$$= \varphi_{p,1}\, S_{(t-1)} + \varphi_{p,2}\, S_{(t-2)} + \ldots + \varphi_{p,p}\, S_{(t-P)} + a_{(t)}$$

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

Selection of order of the AR (p) model:

Residual variance method should be used for selection of model order.

In this method, residual variance $S_Z^2(p)$ was calculated using the following equation for different orders.

$$S_Z^2(p) = \frac{1}{n-2p-1} S(\mu(\alpha_1, \alpha_2, \ldots, \alpha_p) \quad (7)$$

where,

S (m, a_1, a_2.........a_p) is known as residual sum of squares and

a_1, a_2.........a_p parameter of the corresponding model,

The minimum value of $S^2_z(p)$ suggests an approximate order of the autoregressive model to be used further.

Based on the above procedure, order of the AR model was selected and the model was tried with time series data.

Estimation of autoregressive parameters:

The parameter estimation deals with the estimation of the autoregressive parameters of equation (6). These parameters can be expressed in terms of serial correlation coefficient, as Yule-Walker equations. The general recursive formulae for estimating these parameters ($\phi_{p,k}$), where suffix p and k indicate the order and the number of parameter of the order in AR (p) model, respectively may be written as:

$$\varphi_{p,k} = \left[\frac{r_p - \sum_{k=1}^{p-1} (\varphi_{p-1,k})(r_{p-k})}{1 - \sum_{k=1}^{p-1} (\varphi_{p-1,k})(r_k)} \right] \quad (8)$$

or

$$\varphi_{p,k} = \varphi_{p-1,k} - \varphi_{p,p} \cdot \varphi_{p-1,p-k} \; ; k = 1,2,3....p - 1 \quad (9)$$

In equation (8) r_k is the autocorrelation coefficient, autocorrelation regression of the series for K and was computed, for any series $Y_{(t)}$ at any lag, l, as follows:

$$r_l = \sum_{t=1}^{N-1} [Y_{(t)} - \overline{Y}_{(t)}][Y_{(t+1)} - \overline{Y}_{(t)}] / \sum_{t=1}^{N} [Y_{(t)} - \overline{Y}_{(t)}]^2 = C_1 / C_o \quad (10)$$

After estimating the AR parameters $f_{p,k}$, $S_{(t)}$ was calculated by using equation (6). A computer was used for computation of model parameters and stochastic components.

The sum of the periodic and stochastic component forms the generated value of the observed data. The difference was termed as residuals which were tested to check the adequacy of the formulated model.

Diagnostic checking of the model:

Diagnostic checking concerns the verification for the adequacy of the fitted model. The examination of the autocorrelation structure of the residuals provides a powerful way of diagnostic checking.

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.

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

Following tests were conducted to check the adequacy of the model

EXPERIMENTAL FINDINGS AND ANALYSIS

For testing the statistical characteristics of maize evapotranspiration series, eight years data of pan evaporation was taken. Relationship between measured evapotranspiration and pan evaporation was developed. Using this relationship and daily pan evaporation values, maize evapotranspiration values of eight years (1998-2005) were generated. The statistical characteristics of the evapotranspiration series were estimated. Maize evapotranspiration data show that values vary from 1.5375 mm/day to 4.9375 mm/day. Mean value of maize evapotranspiration was found to be 3.52 mm day^{-1}. There is large variability among the values of maize evapotranspiration of different years. The variation may be attributed towards the natural changes in yearly climate. The estimated variance indicates that the coefficient of variation ranges from 0.2077 to 1.270. This signifies the importance of variability of maize evapotranspiration series. Since the values of variance significantly different from zero, it confirms that maize evapotranspiration is mutually dependent.

Serial correlation coefficient:

The lag one serial correlation coefficient of observed series was calculated by using Equation (10) and was found to 0.837. The respective confidence limits were estimated as 0.083 and –0.083 (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 maize evapotranspiration process is mutually dependent. From the analysis of coefficient of variation and serial correlation, it is confirmed that maize evapotranspiration process is time variant and not an independent one. Thus the maize evapotranspiration series may be modelled on stochastic theory. The mutual dependence of the observed maize evapotranspiration series was also confirmed by the correlogram (Fig. 1).

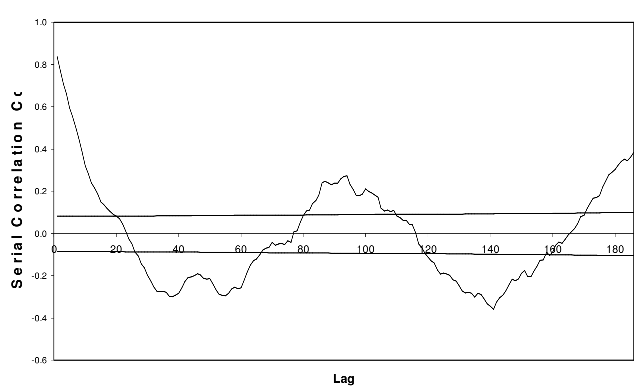

Fig. 1 : Correlogram of observed series of maize evapotranspiration for eight years (1998-2005)

Trend component:

For identification of trend components, maize evapotranspiration series was used. The maize evapotranspiration series was obtained by transforming the eight years annual series. For detection of trend the hypothesis of no trend was made and value of test statistic (z) was calculated by Turning Point test and Kendall's Rank Correlation test. The calculated values of test statistic (z) are -0.77 and 0.56 for Turning Point test and Kendall's Correlation test, respectively. The estimated value of test statistics (z) obtained for Turning Point test, and Kendall's Rank Correlation test was within the 1 per cent level of significance. Hence, the hypothesis of no trend was accepted.

Further from the Turning Point test total numbers of turning points in series were found to be 2. This indicates that the evapotranspiration series is random. From the above analysis it is confirmed that the trend component in evapotranspiration time series is absent and the observed series may be treated as trend free series.

Periodic component:

To confirm the presence of periodic component in maize evapotranspiration series a correlogram was drawn. A correlogram is a graphical representation of serial correlation coefficients (r_l) as function of lag l in which the values of r_l are plotted against respective value of l. The correlogram of the observed series is shown in Fig. 1. The resulting oscillating shape of the correlogram confirms the presence of periodic component in the monthly evapotranspiration.

Further, the correlogram has peaks at legs equal to 92 and at other multiples of it. The time span of periodicity was taken as 92 for use in harmonic analysis of periodic component.

Determination of significant harmonics:

For representing the periodic component of the maize evapotranspiration 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 can be ignored.

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 graph was drawn between P_i and number of harmonics, called the cumulative periodogram. 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 maize evapotranspiration series was made and the plot of Pi against i is shown in Fig. 2. From the cumulative periodogram in the maize evapotranspiration series, it can be observed that the first two 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.

For the first two harmonics the values of Fourier coefficients (A_1, A_2 and B_1, B_2) were found to be -1.372, -0.195, -0.373, 0.287, 0.242, and - 0.055, respectively. With these coefficients and using equation (2) the periodic component, $P_{(t)}$ resulting from periodic deterministic process may be mathematically expressed as:

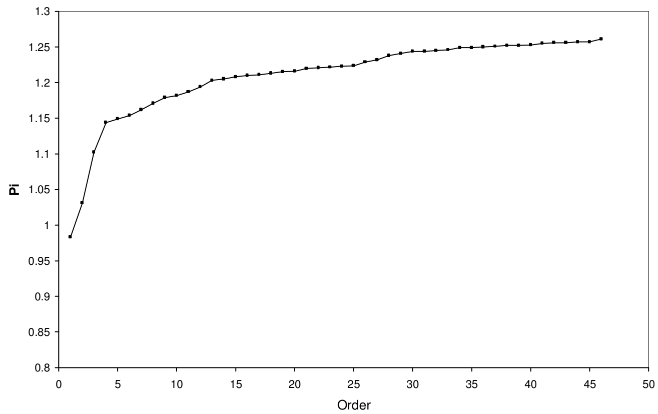

Fig. 2 : Cumulative periodogram of maize evapotranspiration

$$P_{(t)} = 3.52 - 1.372 \text{ Cos } (2pt/P) + 0.287 \text{ Sin } (2pt/P) - 0.195 \text{ Cos } (4pt/P) + 0.242 \text{ Sin } (4pt/P) -0.373 \text{ Cos } (6pt/P) - 0.055 \text{ Sin } (6pt/P) \quad ...(12)$$

The deterministic cycle component, $P_{(t)}$ was computed by using equation (12) for all the values of t (t_{max}= 244). After estimating the periodic component it was removed from, historical time series by subtracting the periodic component from historical time series. This process obtained a new stationery series, S_t, resulting from stochastic non-deterministic process.

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).

The periodic component was removed from the historical series. The rest of the data were analyzed to obtain non-deterministic stochastic component by fitting the autoregressive of stochastic modeling.

Estimation of autoregressive parameters

The autoregressive parameters were estimated by using equations (8), (9) and (10). The estimated values of auto covariance function and SCC of different lags (l_{max}= 186) indicate the linear dependence between each lag. A study reveals that the values of SCC are significantly different from zero which confirms the dependence of present values and past values. In other words, it may be concluded that the past and present values are highly inter correlated.

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. The minimum residual, variance was obtained for order three. The values of residual variance after 3rd order showed no definite trend.

Using equation (6) and the estimated autoregression coefficients the stochastic component of the maize evapotranspiration time series may be expressed as:

$$\mathbf{S = 0.620S_{t-1} + 0.192\ S_{t-2} + 0.014\ S_{t-3} + 0.064\ S_{t-4}} \quad \mathbf{(13)}$$

The non-deterministic stochastic component was estimated by using equation (13) for all the values ($t_{max} = 2678$).

The residual series of stochastic component:

The residual series, $a_{(t)}$ which is random independent part of stochastic component was obtained 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 0.0 and SD 0.5 mm day^{-1}). The values of statistical measures are presented in 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 evapotranspiration series was found to be a trend free series that developed model describes the periodic-stochastic behaviour of the series.

The developed model 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:

$$\mathbf{X(t) = 3.52 - 1.372\ Cos\ (2pt/P) + 0.287\ Sin\ (2pt/P) - 0.195\ Cos\ (4pt/P) + 0.242\ Sin\ (4pt/P) - 0.373\ Cos\ (6pt/P) - 0.055\ Sin\ (6pt/P) + .620\ S_{t-1} + 0.192\ S_{t-2} + 0.014\ S_{t-3} + 0.064\ S_{t-4} + a_t} \quad ...(14)$$

The first seven terms in the formulated model represented by equation (14) constitute the deterministic part of the evapotranspiration time series. The eighth, ninth and tenth term represent the dependent stochastic component of the model where the current value of S_t 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 evapotranspiration series was generated for all the values (t = 744).

Diagnostic checking of maize evapotranspiration 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 maize evapotranspiration.

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 0.9839, 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 maize evapotranspiration series.

Serial correlation analysis

The serial correlation coefficients (SCC) for lags l (l= 1, 2, 3, ….186) were computed with the help of equation (10). The values of SCC against respective lags were then plotted to obtain a correlogram. The resulting correlogram is shown in Fig. 3 with confidence limit at 1 per cent level. The correlogram is almost completely contained within the confidence limits at 1 per cent level. Hence, it may be treated to be non significant. Further the correlogram shows that at 1 per cent level the coefficients are within the limits. This again confirms that the residual series may be treated to be non significant.

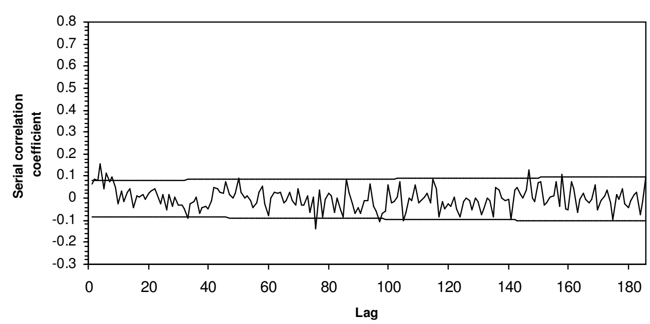

Fig. 3 : Correlogram of residual series of daily maize evapotranspiration for eight years (1998-2005)

The residual series has a mean value of 0.0001 (zero) and the variance of 0.06. This leads to the conclusion that the residuals are independent normally distributed. Further, it also confirms the randomness of the residuals.

Table 1 : Statistical parameters of the observed, generated and residual series of evapotranspiration

Series	Mean mm day^{-1}	Standard deviation mm day^{-1}	Variance
Historical series	3.523306	1.756899	3.086694
Generated series	3.523465	1.577098	2.487237
Residual series	-0.00016	0.24497	0.06001

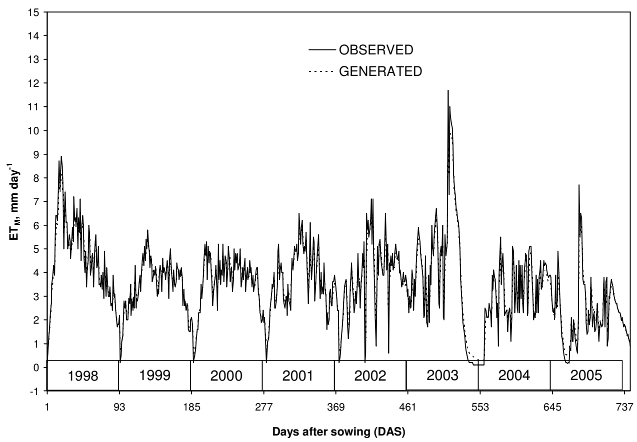

Fig. 4 : Variation of generated (ET_{MG}) and observed (ET_{MO}) mean daily maize evapotranspiration for eight years at Banswara

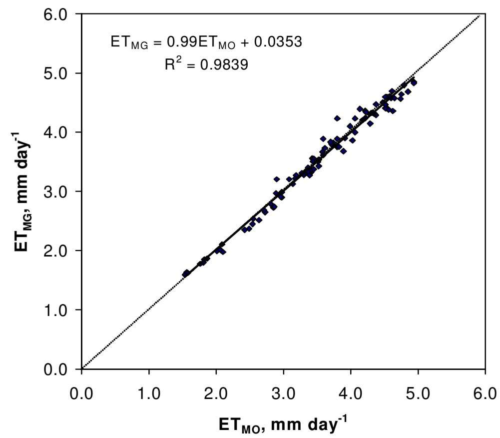

Fig. 5 : Relationship between generated and observed mean daily maize evapotranspiration for eight years

Validation of stochastic model of maize evapotranspiration:

Validation of generated maize evapotranspiration series developed by stochastic model (equation 14) was done by comparison of generated maize evapotranspiration series and measured maize evapotranspiration series. Validation of generated 8 years maize evapotranspiration series was made with 8 years mean measured evapotranspiration series (Fig. 4). The relationship is shown in Fig. 5. The correlation coefficient between generated maize evapotranspiration series and measured maize evapotranspiration was found to be 0.99. The correlation was tested by t test and found to be highly significant at 1 per cent level. The standard error (0.13 mm day^{-1}) is quite low. The mean of the generated maize evapotranspiration was found to be 3.5234 mm day^{-1}. Mean of the measured maize evapotranspiration series was found to be 3.5233 mm day^{-1}. The regression equation is very near to 1:1 line. Therefore, equation (14) can be used for future prediction of maize evapotranspiration.

List of abbreviations:

$\overline{Y}_{(r)}$ = mean of the series, $Y_{(t)}$

$\phi_{P,K}$ = autoregressive model parameters, K = 1,2 ...,p.

$a_{(t)}$ = independent random number

A_K and B_K = Fourier coefficients.

C_l = autocovariance function at lag, l, 1= 0, 1, ...p

K = number of significant harmonics

N = total number of discrete values of $X_{(t)}$

N = number of observation points and

p = base period

$P_{(t)}$ = periodic component

$S_{(t)}$ = stochastic component, including dependent and independent parts.

$T_{(t)}$ = trend component, t = 1,2 ....N

REFERENCES

Kottegoda, N.T. (1980). *Stochastic water resource technology*. The Macmillan Press Ltd., London. 384 pp.

Nirmal, S.V. (2011). Modelling of evapotranspiration for maize and gram crops for Banswara. Ph.D. Thesis, Department of Soil and Water Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, UDAIPUR (RAJASTHAN) INDIA

****