Journal of Agrometeorology 6 (2) : 153 - 161 (Dec 2004)

Stochastic modeling of evapotranspiration*

R. SUBBAIAH

College of Agricultural Engineering and Technology, Junagadh Agricultural University, Junagadh 362 001

ABSTRACT

The stochastic structure of weekly evapotranspiration time series estimated from the Penman equation is analyzed using an additive model. The turning point test and Kendall's rank correlation test are applied for detecting the trend. Correlogram technique is used to detect the periodicity, which is then analyzed by Fourier series method. Significant harmonics were identified . The series is then tested for stationarity and the dependent part of the stochastic is found to be well expressed by the second order auto regressive model. As a result, the model superposes a periodic deterministic process and a stochastic component. The developed periodic–stochastic model may be used for representing the time-based structure of the evapotranspiration time series.

Key Words: Stochastic, evapotranspiration, stationary, periodic.

Quantification of potential evapotranspiration (PET) is required for plant production, management of water resources, environmental assessment, irrigation scheduling and in designing water storage and distribution systems. Most of the research workers considered evapotranspiration (ET) either to be deterministic or probabilistic in nature. While the former methods do not consider the random effects of various input parameters, the later methods employ the concept of probability to the extent that the time-based characteristics of ET are ignored. ET is stochastic in nature because it is affected by climatological parameters i.e. stochastic climatic variations are transformed to become stochastic component of ET. Hence, the ET needs to be computed by considering both the deterministic and stochastic part of the process. Stochastic analysis of ET time series will provide a mathematical model that will account for the deterministic and stochastic parts and will also reflect the weekly variations of crop water demands.

During the past years, many investigators (Quimpo and Rafel, 1968; Dyer, 1977; Spolia *et al* 1980; Srikanthan and McMohan, 1982; Raghuvanshi and Walender, 2000) have analyzed the time series of rainfall and stream flow and

*Paper presented in the National Seminar on "Agrometeorological Research for Sustainable Agricultural Production", held at G.A.U., Anand during 27-28 September 2001.

developed the auto regressive, trigonometric regression and other forms of the stochastic models for generation of the data. But no efforts were made to study the structure of the ET. With this in view the present study was under taken with the following specific objectives: 1) to test the ET time series for stochasticity; 2) to identify and remove the trend and periodic components; 3) to study the structure of dependent stochastic component and express the time dependence by an appropriate stochastic model; and 4) to diagnostically check for the independent stochastic component.

MATERIAL AND METHODS

Penman's method is used to estimate the reference ET. The details of this method are presented in Doorenbos and Pruitt (1977). Wright (1982), Richard and Pruitt (1991), Cuenca (1989) and Snyder (1992) developed regression equations for calculating the correction factors used in the above approaches. The Penman's ET series were used to develop the stochastic model. A general additive model is used to describe the ET time series, and is given by:

$$X_t = T_t + P_t + S_t \tag{1}$$

Where T_t =the trend component at time, 't', t = 1,2,....N; P_t = the periodic component; S_t= the stochastic component having dependent and independent parts; and N= the number of data points. Eq. 1 was systematically identified and its components were removed. The adopted procedures are described in the following subsection.

Trend component

The trend of the data was identified for the weekly ET values (Z_i, i=1,2,3....n, n is the number of weeks) using the turning point test and Kendall's rank correlation test. The details of these tests are presented in Clarke (1973). If the trend is present, then T_t was removed by regression. After removing the trend, a trend free series was obtained as:

$$Y_t = X_t - T_t = P_t + S_t \tag{2}$$

Periodic component

The Y_t can be expressed in Fourier form as follows:

$$Y_t = A_0 + [\sum_{K=1}^{P/2} A_k \operatorname{Cos}(2 \pi k t/P) + B_k \sin(2 \pi k t/P)] \tag{3}$$

in which P = time span of periodicity; k = number of harmonics; $1 \le k \le P/2$; M = number of significant harmonics; $1 \le M \le P/2$; and N = number of data points. The Fourier coefficients A_k and B_k of eq. 3 were computed by the following formulae:

$$A_k = (2/N) \sum_{t=1}^{N} [Y_t \operatorname{Cos}(2 \pi k t/p)] \tag{4a}$$

$$B_k = (2/N) \sum_{t=1}^{N} [Y_t \sin(2 \pi k t/p)] \tag{4b}$$

$$A_0 = (1/N) \sum_{t=1}^{N} [Y_t] \tag{4c}$$

in which N = the number of data

points. Analysis of variance test was utilized to determine the number of significant harmonics. For this test, periodic means m_r were computed and are expressed in Fourier form as follows:

$$m_r = m_0 + [\sum_{i=1}^{N} [\alpha_k \, Sin\,(2 \pi k t / P) + \beta_k \, Cos\,(2 \pi k t / P)] \quad (5)$$

$$m_0 = \sum_{r=1}^{p} [m_r / P] \quad (6a)$$

$$\alpha_k = (2 / P) \sum_{r=1}^{p} [m_r \, Sin\, (2 \pi k r / p)] \quad (6b)$$

$$\beta_k = (2 / P) \sum_{r=1}^{p} [m_r \, Cos\,(2 \pi k r / p)] \quad (6c)$$

In this test, the null hypothesis was that the variance explained by harmonic K, which was (N/2) ($\alpha_k^2 + \beta_k^2$) is zero. Computations were made to test the α_k and β_k values for k = 8,7...1 in order to obtain the F-ratio, which was then compared with its table value at 0.01 level of significance. Thus the number of significant harmonics was obtained and P_t was computed, using eq. 3 for these harmonics only. The computed P_t was removed from Y_t, which leaves only the stochastic component, S_t, to be analyzed further.

Stochastic component

It was assumed that the value of S_t, at time t was the combined effect of the weighted sum of the past values (whole year) so that the dependent part of S_t may be represented by the equation:

$$S_t = [\sum_{k=1}^{\infty} \phi_{p,k} \, S_t + a_t] \quad (7)$$

in which $\phi_{p,k}$ = the autoregressive parameter; k= the number of parameters, k =1,2, ...P; P = the order of the model; and a_t = the independent and normally distributed error variable. Because of the diminishing effect of the past values on the present, the upper limit of Eq. 7 may be made finite, say p, resulting in a finite order Markov model:

$$S_t = [\sum_{k=1}^{p} \phi_{p,k} \, S_t + a_t] \quad (8)$$

$$\phi_{p,k} = [\phi_{p-1,k} - \phi_{p,p} \, (\phi_{p-1,\, p-k})] \quad (9)$$

$$\phi_{p,p} = [r_p - \sum_{k=1}^{p-1} \{ (\phi_{p-1,k}) r_{p-k} \} / 1 - \sum_{k=1}^{p-1} \{ (\phi_{p-1,\, p-k}) \gamma_k \}] \quad (10)$$

$$r_p = C_p / C_0 \quad (11)$$

$$C_p = E [(S_t - \mu) \, (S_{t+p} - \mu] \quad \text{where } \mu = E(S_t) \quad (12)$$

The model represented by eq. 7 is known as the autoregressive model of order p, AR(p). r_p = autocorrelation coefficient; C_p and C_o = auto covariance function at lag p and lag 0; and E(.) = mathematical expectation. The dependent part of S_t was obtained by eq. 7 and then was removed, leaving the independent part as follows:

$$a_t = [S_t - \Sigma(\phi_{p,k} \, S_{t-k})] \quad (13)$$

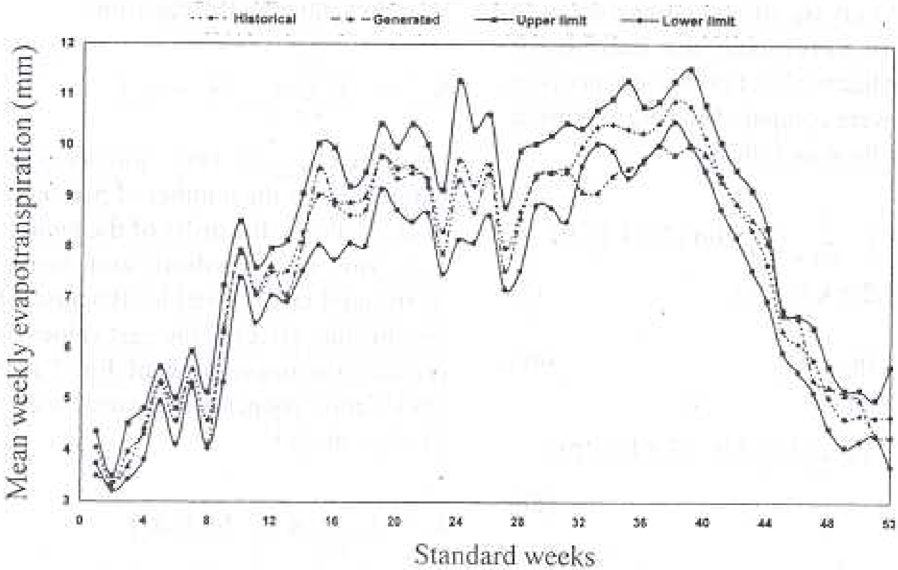

Fig.1 : Comparison between historical and generated evapotranspiration series

The a_t series is called the residual series. The fitting procedure on this model involves two stages (Box and Jenkins, 1976): (1) Selection of the model order, p; and (2) estimation of autoregressive coefficients, $\phi_{p,k}$. For selection of order p, Residual Variance Method (Shibata, 1985) Porte Menteau Statistic (proposed by Box and Pierce, 1970 and later modified by Ljung and Box, 1980), Akaike Criterion (Akaike, 1974) was adopted.

Diagnostic checking

Diagnostic checking means statistically verifying the adequacy of the formulated model. For this checking, the residual series was examined for any lack of randomness. Auto-correlation coefficients of residual series for lag-1 (l-60) were computed and were drawn against lag with 95 % tolerance limits. If the correlogram thus obtained is well within the limits, then it can be inferred that residuals are normally distributed with zero mean and Var (1/l). Various other test details were also done to confirm the randomness of the residuals, which is the condition for accepting the formulated auto-regressive model.

Data

The climatic data was collected from the meteorological observatory of Gujarat Agricultural University Junagadh for the past 21 years (1980-2000). The ET was estimated for each years using the Penman method. The mathematical procedure described above is utilized to investigate the structure of the time series of ET for Junagadh.

Table 1: Statistical properties of weekly Penman evapotranspiration

Week	Turning point test	Kendall rank Correlation test	Auto correlation coefficient					Mean (mm)	Standard deviation	Serial correlation Coefficient
			r_1	r_2	r_3	r_4	r_5			
01	-1.93	1.878	0.25	-0.37	-0.29	-0.16	0.13	4.050	0.606	0.314
02	-1.10	1.633	0.01	-0.51	0.11	0.26	-0.33	3.351	0.280	0.254
03	-1.10	-0.889	0.03	-0.35	0.27	0.12	-0.34	3.971	1.068	0.128
04	-1.10	-0.355	0.15	-0.30	-0.14	-0.18	-0.23	4.313	0.984	-0147
05	0.55	-0.914	-0.16	-0.33	0.04	-0.33	0.13	5.286	0.654	-0.315
06	-1.10	-2.155	-0.08	-0.16	-0.04	0.13	-0.11	4.563	0.902	0.211
07	-1.93	-0.635	0.25	-0.45	-0.57	-0.17	0.34	5.575	0.689	0.322
08	-1.10	1.071	0.09	-0.27	-0.47	0.09	-0.06	4.585	1.072	0.151
09	0.55	0.666	-0.15	-0.05	-0.09	0.15	-0.09	6.288	1.901	0.231
10	-1.10	-1.824	0.10	-0.08	0.06	-0.11	-0.13	7.946	1.098	0.287
11	-1.10	-1.655	0.41	0.03	-0.21	-0.19	-0.14	7.044	1.081	0.241
12	1.38	-0.988	-0.13	-0.14	-0.33	-0.03	0.02	7.504	0.920	0.264
13	1.38	-0.873	-0.20	0.07	0.12	0.19	-0.28	7.527	1.188	0.341
14	0.55	0.901	-0.65	0.24	-0.18	0.30	-0.19	8.295	1.503	0.136
15	-1.10	0.971	-0.33	-0.23	0.16	-0.04	-0.26	9.032	1.961	0.158
16	-0.27	0.853	-0.33	-0.55	0.43	0.15	-0.28	8.847	2.177	0.147
17	-1.10	-0.333	0.41	-0.16	-0.03	-0.07	-0.42	8.636	1.106	0.198
18	0.55	-1.118	0.25	0.00	-0.07	-0.21	-0.32	8.746	1.432	0.245
19	-1.10	-1.104	-0.12	-0.39	0.06	-0.18	-0.07	9.780	1.259	0.311
20	-0.27	-0.687	0.17	-0.35	-0.36	-0.25	0.03	9.353	1.215	0.412
21	-1.93	1.071	0.08	-0.42	-0.35	-0.06	0.36	9.460	1.885	0.428
22	0.55	0.348	-0.02	-0.16	-0.09	-0.19	-0.03	9.351	1.317	0.437
23	0.55	-0.494	-0.08	-0.45	0.00	0.16	-0.15	8.306	1.634	0.341
24	-1.10	-0.828	0.12	-0.39	0.02	0.32	-0.24	9.711	3.092	0.321
25	-1.10	1.082	0.14	-0.05	0.01	-0.23	-0.28	9.206	2.190	0.311
26	0.55	1.028	-0.06	-0.22	-0.49	0.06	0.12	9.625	1.955	0.258
27	-0.27	1.113	-0.09	-0.25	-0.05	0.00	0.00	7.992	1.614	0.118
28	-0.27	0.821	-0.10	-0.47	-0.11	0.17	0.13	8.716	2.323	0.195
29	-0.27	-0.666	-0.25	0.01	-0.20	-0.21	0.10	9.431	1.192	0.162
30	-1.93	-0.961	0.36	-0.43	-0.33	0.16	0.25	9.488	1.394	0.135
31	-0.27	-1.895	0.18	0.04	-0.37	0.20	0.03	9.478	1.902	0.104
32	-1.93	0.577	0.22	-0.20	-0.30	0.04	-0.25	9.967	0.735	0.162
33	-1.10	-1.136	0.29	-0.31	-0.36	-0.06	0.14	0.362	0.649	0.111
34	-1.93	-0.003	0.44	-0.15	-0.22	0.14	0.18	10.411	1.004	0254
35	0.55	-0.483	-0.07	-0.08	-0.20	0.39	-0.15	10.322	1.874	0.333
36	-0.27	0.247	-0.03	-0.28	-0.20	0.18	0.21	10.253	1.031	0.218
37	-0.27	1.077	0.24	0.20	-0.14	0.19	-0.13	10.444	0.835	0.128
38	-1.93	2.224	0.25	-0.28	-0.18	0.19	-0.15	10.876	0.780	0.364
39	-1.10	2.554	0.24	-0.25	0.06	0.23	-0.09	10.775	1.468	0.374
40	-0.27	1.997	0.16	-0.02	-0.03	0.14	-0.10	10.155	1.265	0.298
41	-0.27	0.003	-0.09	-0.41	0.10	0.36	-0.23	9.415	1.288	-0.118
42	-0.27	1.895	0.15	-0.13	0.03	0.17	-0.03	8.865	1.306	0.185
43	-1.10	2.389	-0.33	-0.16	0.07	0.14	-0.03	8.352	1.541	0.411
44	-1.93	-1.565	0.11	-0.54	-0.43	0.07	0.29	7.674	1.430	0.354
45	-0.27	-1.111	-0.04	-0.17	-0.06	0.25	-0.31	6.354	0.805	0.195
46	-1.93	-1.214	0.25	-0.32	-0.14	-0.08	-0.14	6.109	1.086	0.187
47	-1.10	-1.381	0.16	-0.33	-0.21	0.25	-0.11	5.785	1.252	0.116
48	-1.10	-1.088	0.15	-0.75	-0.20	0.46	0.10	5.048	1.202	0.112
49	-1.93	2.371	0.46	0.06	-0.02	-0.13	-0.10	4.639	1.105	-0.109
50	1.38	0.577	0.41	0.02	-0.18	-0.11	-0.34	4.700	0.985	-0.184
51	-0.27	0.628	0.18	0.18	-0.32	-0.15	-0.29	4.660	0.699	-0.312
52	1.38	0.885	-0.46	-0.02	-0.05	0.20	-0.21	4.680	1.925	0.245

Table 3: ARMA Model estimates for evapotranspiration estimated from Penman approach

Parameter	ARMA (p, q) Model order			
	(1,0)	(2,0)	(3,0)	(4,0)
ϕ_1	0.503	0.488	0.486	0.487
ϕ_2		0.029	- 0.009	- 0.009
ϕ_3			0.077	0.081
ϕ_4				- 0.008
Residual variance (Unbiased)	0.791	0.775	0.786	0.787
Porte Manteau Statistic (χ^2 95)	24.0	12.1	22.2	21.1
Akaike Information Statistic	- 420	- 411	- 426	- 422

Table 4: Autocorrelation coefficients of residuals

S No	Autocorrelation coefficients of residuals (a_t)			
	Lag 1-15	Lag 16-30	Lag 31-45	Lag 46-60
1	-0.004	-0094	-0.066	-0.065
2	0.221	0.116	0.132	0.018
3	-0.054	0.028	0.056	-0.011
4	0.221	0.158	0.156	0.016
5	0.105	0.118	0.260	0.295
6	-0.233	-0.181	-0.093	0.044
7	-0.087	-0.018	0.045	-0.021
8	0.065	0.051	0.089	0.113
9	0.055	-0.031	0.014	0.011
10	0.155	-0.103	0.008	-0.027
11	0.105	0.117	0.261	0.285
12	-0.068	-0.207	-0.156	-0.041
13	-0.191	-0.225	0.288	-0.108
14	0.221	0.156	0.016	0.166
15	-0.227	-0.193	-0.122	0.300

obtained were removed from the original time series in order to get a new stationary series.

The stochastic component was analyzed by fitting the auto regressive process to the series. The order of the model was determined by the procedure explained earlier. The auto regression coefficients obtained in this study were confirmed for their stationarity. The residual variance, Porte Manteau Statistic and Akaike Information Statistic (AIC) were found to be 0.775, 12.1 and −411•for the second order Markov model. Hence, S_t can be approximated by the second order Markov model. As the ET derived from various approaches is a trend free series, the developed model would describe the periodic stochastic behaviour of the original series. It is a superposition of a harmonic deterministic process and second order Markov model.

The formulated model was subjected to various checks to test its adequacy for representing the time dependent structure of the ET series derived from various approaches. The autocorrelation coefficients of the residuals for various lags (Table 4) were computed. After comparing the correlogram with 95 % confidence limits, the results revealed that almost all the coefficients are small and hence could be treated as non-significant. The autocorrelation coefficient has a mean and variance values of 0.0113 and 0.0131 respectively which approximately equal to Var (1/60) = 0.0167. This leads to the conclusion that the residuals are independent and normally distributed. The turning point test has also confirmed the randomness of the residuals, corroborating the model. The ET values are generated by the formulated model and are plotted with the corresponding observed values. The plot indicated closeness of the values and thereby reflects the appropriateness of the formulated ET model. Therefore the model may be employed to generate the weekly ET values, which can be used in planning and operation of irrigation projects.

CONCLUSION

The objective of this study on the stochastic analysis of ET is to formulate a mathematical model of the stochastic ET. The study revealed that the developed model is feasible. It was found that the weekly ET series are trend free and periodic and are stochastic in nature. The developed model superposes a periodic-deterministic process and a stochastic component. The deterministic portion has been analyzed using the Fourier series. The time dependence of the stochastic portion was well approximated by the second order autoregressive model. The removal of this dependence has lead to a series of independent normal random variables. The significance of this study is to show that the past records of the data provide valuable information for determining the basic time dependent structure of ET series.

REFERENCES

Akaike, H. 1974. A new look at statistical identification. *IEE Trans. Auto.Control, AC* 19 (2): 716-722.

Box, G.E.P. and Pierce, D.A. 1970. Distribution of residual autocorrelation in autoregressive integrated moving average time series models. *J. American Statis. Assoc.*, 70 (1): 70-79.

Box, G.E.P. and Jenkins, G.M. 1976. Time series analysis: forecasting and control

Holden day, Sanfrancisco, USA., pp 600.

Clarke, R.T.1973. Mathematical models in Hydrology, Irrigation and Drainage Paper, FAO, Rome, pp 282.

Cuenca, R.H.1989. Irrigation system design an engineering approach, Prentice Hall, Engle Wood Cliffs, New Jersey, pp 133.

Doorenbos, J. and Pruitt, W.O. 1977. Crop water requirements, Irrigation and Drainage Paper, FAO, Rome, pp 305.

Dyer, T. G. J. 1977. On the application of some stochastic models to precipitation forecasting. *Quart. J. Roy. Meteorol.Soc.*, 103 (1) : 177-89.

Ljung, G.M. and Box, G.E.P. 1978. On a measure of lack of fit in time series models. *Biometrica*, 65 (2): 297-304.

Quimpo, R. and Rafel, H.G. 1968. Stochastic analysis of Daily river flows. *J. Hydraulics Division, ASCE*, 94 (1): 43-50.

Raghuvanshi, N.S. and Wallender, W.W. 2000. Stochastic modeling of evapotranspiration. *Intern. J. Agric. Engg.*, 2 (1) :25-37.

Richard, G.A. and Pruitt, W.O. 1991. FAO 24 Reference evapotranspiration factors. *J. Irrig. Drain. Div., ASCE*, 117 (5): 758-773.

Shibata, R.1985. Various model selection techniques in time series analysis in E J Hannan *et. al.* Hand book of Statistics, 5:179-187.

Snyder, R.L. 1992. Equation for evaporation pan to evapotranspiration conversion. *J. Irrig. Drain. Div., ASCE*, 118 (6): 977-980.

Spolia, S.K., Chander, S. and O-Connor, K.M. 1980. An autocorrelation approach for parameter estimation of fractional order equal root autoregressive models using hyper geometric function. *J. Hydraulics Div., ASCE*, 47 (1) :1-17.

Srikantan, R and MacMohan, T.A. 1982. Stochastic generation of monthly stream flows. *J. Hydraulics Div., ASCE*, 108 (3): 419-428.

Wright, J.L.1982. New ET crop coefficients. *J. Irrig. Drain. Div., ASCE*, 108 (1): 57-74.