

蒸散量变化的随机模型*

龚元石 李保国

(中国农业大学土壤和水科学系,北京 100094)

摘要: 根据参考作物蒸散量时间序列的随机变化特征,本文探讨了建立蒸散量变化的自回归滑动平均(ARMA(p, q))模型的方法。应用矩估计法对模型作初步估计,用最小二乘法对ARMA(p, q)模型参数作精估计,模型最佳阶数判别采用AIC准则。结果表明,蒸散量的随机变化特征可用ARMA(1,1)模型描述。最后,根据建立的模型对蒸散量进行了预测。

关键词: 参考作物蒸散量; 随机模型

中图分类号: S161.4

蒸散量(这里指参考作物蒸散量)是在一定条件下土壤蒸发和作物蒸腾量的总和,它与众多气象因子如温度、湿度、风速和日照等有关,具有时空分布不均的特点,尤其在时域上,它既有年、月、日内的周期变化趋势,又有其相对应的随机变化特征。这种随机变化特征很难用确定性的数学模型加以描述。

无论是研究流域范围内的水分循环,还是农田尺度上的水分平衡,蒸散量都作为一个主要的分量起着重要的作用。因此,精确估算或预测蒸散量具有重要的科学价值。特别是在水资源日益紧缺的今天,合理用水,节约用水已成为全人类共同关注的问题。在农田水分管理方面,正确估算土壤水分和制定合理的灌水决策是实现节水农业的重要措施之一,而土壤水分的变化与蒸散量的变化是密切相关的。在已经建立的土壤水动力学模型^[1]以及土壤水分平衡模型^[2]中,蒸散量一般作为模型的边界条件或者水分平衡分量的已知项输入,实际上,由于蒸散量在时域上变化的不确定性,当上述模型用以预测土壤水分时就变得相当困难。显然,如何根据蒸散量的随机变化特征,建立蒸散量变化的随机模型,对提高已经建立的水均衡模型的预测能力,将具有重要的意义。

已有研究表明^[3],蒸散量在时域上的变化可用自回归滑动平均ARMA(p, q)模型来描述,Aboitiz(1986)研究了蒸散量的随机变化规律,并根据建立的随机模型结合土壤水分平衡模型预测了土壤水分变化。Or和Hanks(1993)则根据建立的随机模型结合作物产量模型用于优化农田灌溉决策^[4]。

本文根据北京地区参考作物蒸散量时间序列的随机变化特征,探讨建立蒸散量变化的自回归滑动平均模型的方法,并利用建立的模型对蒸散量进行预测。

1 蒸散量变化的随机模型

1.1 自回归滑动平均模型ARMA(p, q) 根据蒸散量的随机变化特征,其随时间变化过程可用一般的自回归滑动平均模型进行描述^[3]:

收稿日期: 1995-03-20

*属国家自然科学基金资助项目

$$ET_t = \varphi_1 ET_{t-1} + \dots + \varphi_p ET_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (1)$$

式中:

φ_i : 自回归(AR)系数($i=1, 2, \dots, p$)

p : 自回归阶数

θ_i : 滑动平均(MA)系数($i=1, 2, \dots, q$)

q : 滑动平均阶数

ET_t : 标准化后的蒸散量

ε_t : 零均值正态分布的随机项, 方差为 σ_ε^2

建立 ARMA(p, q)模型的过程实际上是要确定(1)式中模型系数(φ_i, θ_i)和阶数(p, q)以及残差方差 σ_ε^2 , 为了建立模型, 首先必须建立时间序列 $\{ET_t\}$ 的样本。

1.2 蒸散量时间序列的样本处理 根据北京气象站 1979~1988 年 1 月~12 月份的 10 年气象观测资料, 应用 Penman-Monteith 公式^[5]计算历年逐日参考作物蒸散量(ET_p), 并由此计算出逐日 ET_p 的均值($\overline{ET_p}$)及其标准差(STD)(图 1)。从图中看出, 参考作物蒸散量的多年平均值及其标准差的时间序列存在明显的季节变化, 即在年内呈周期变化趋势。为了得到 $\{ET_t\}$ 的平稳时间序列, 首先必须剔除其周期变化特征。周期项可以采用如下的傅里叶级数进行拟合。

$$X_t = a_0 + \sum_{j=1}^k (a_j \cos 2\pi j t / N + b_j \sin 2\pi j t / N) \quad (2)$$

其中, $a_0, a_1, \dots, a_k, b_1, \dots, b_k$ 为傅里叶系数, 分别由下式确定:

$$\left. \begin{aligned} a_0 &= \frac{1}{N} \sum_{j=1}^N X_j \\ a_i &= \frac{2}{N} \sum_{j=1}^N X_j \cos 2\pi i j / N \\ b_i &= \frac{2}{N} \sum_{j=1}^N X_j \sin 2\pi i j / N \end{aligned} \right\} \quad (3)$$

式中 N 为样本个数, k 为有效谐波数, j 为谐波号码, 参考作物蒸散量多年平均值及其标准差时间序列的傅里叶级数公式拟合结果见图 1。图中傅里叶系数只需用 a_0, a_1, a_2 及 b_1, b_2 , 其计算结果见表 1。

表 1 日平均蒸散量及标准差时间序列的傅里叶系数

Table 1 Fourier coefficients for mean daily ET_p and standard deviations series

傅里叶系数 Fourier coefficients	日平均蒸散量 ET_p /mm	标准差 STD /mm
a_0	3.570	1.280
a_1	-2.047	-0.727
a_2	-0.248	-0.026
b_1	0.295	0.086
b_2	-0.411	-0.109

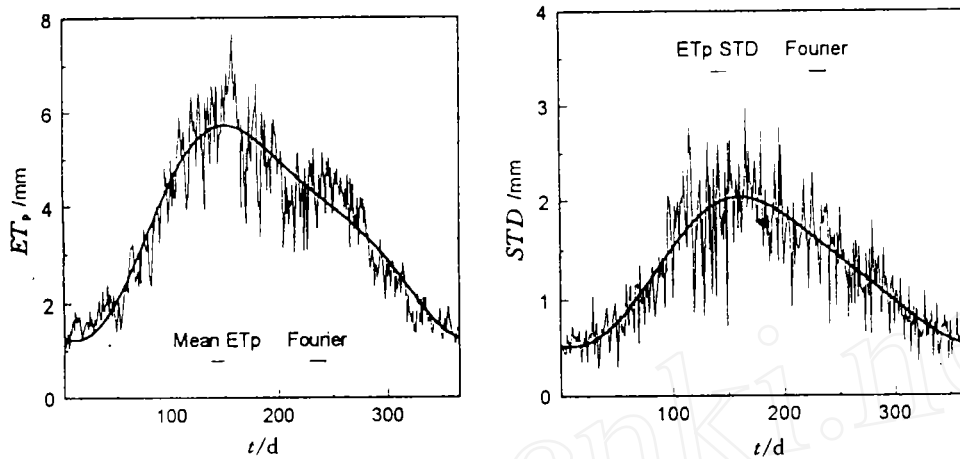


图 1 日平均蒸散量(左)及标准差(右), 季节趋势变化采用两阶谐波傅里叶级数拟合(北京)

Fig. 1 Mean daily ET_p (left) and its standard deviation (right), seasonal trends were fitted with two-harmonic Fourier series (Beijing)

假定把傅里叶级数拟合结果定义为 ET_t , 则剔除周期趋势变化后的蒸散量时间序列离散样本数据 ET_s 为:

$$ET_s = \overline{ET_p} - ET_t \quad (4)$$

进一步对 ET_s 离散数据进行标准化处理:

$$ET_i = \frac{ET_s - \overline{ET_s}}{\sigma_{ET_s}} \quad (5)$$

其中: $\overline{ET_s}$ 为 $\{ET_s\}$ 离散序列的均值, σ_{ET_s} 为 $\{ET_s\}$ 离散序列的标准差。 ET_i 即为所求的标准化随机变量, 其均值为零, 方差为 1。一般认为 $\{ET_i\}$ 一平稳时序过程^[4]。

2 $\{ET_i\}$ 序列随机模型的识别及其参数估计

2.1 模型的初步识别(粗估计) 根据上述确定的样本资料, 应用矩估计法^[6]对模型进行初步识别, 首先计算样本的自相关函数(ρ_r)和偏相关函数(ψ_{kk}), r 和 k 分别为时间滞后期。模型识别的 Box-Jenkins 方法是依据样本自相关函数和偏相关函数的统计特征, 初步判断随机序列 $\{ET_i\}$ 所适合的自回归滑动平均模型, 这种方法可以对 AR(p)或 MA(q)及低阶 AR-MA(p, q)模型进行初步识别。初步识别包括计算样本的均值、自相关函数和偏相关函数。由于 $\{ET_i\}$ 样本经过标准化处理, 所以均值为零。自相关函数和偏相关函数见图 2。

Box-Jenkins 的模型初步识别原则为: 若样本自相关函数 ρ_r 在 $r > q$ 之后截尾, 则判断序列 $\{ET_i\}$ 是 MA(q)模型; 若样本偏相关函数 ψ_{kk} 在 $k > p$ 之后截尾, 则判断序列 $\{ET_i\}$ 是 AR(p)模型; 若 ρ_r 和 ψ_{kk} 都不截尾, 但被负指数型函数所控制(即拖尾), 则应判断其为 AR-MA(p, q)模型, 此时尚不能确定模型阶数。从图 2 中可以看出, 样本的自相关函数 ρ_r 随着 r 增加时逐渐减少, 并趋向于零, 因此, $\{ET_i\}$ 为一平稳序列, 且样本的自相关函数和偏相关函数都具有拖尾的性质, 因此, 可以初步确定 $\{ET_i\}$ 序列为 ARMA(p, q)模型。

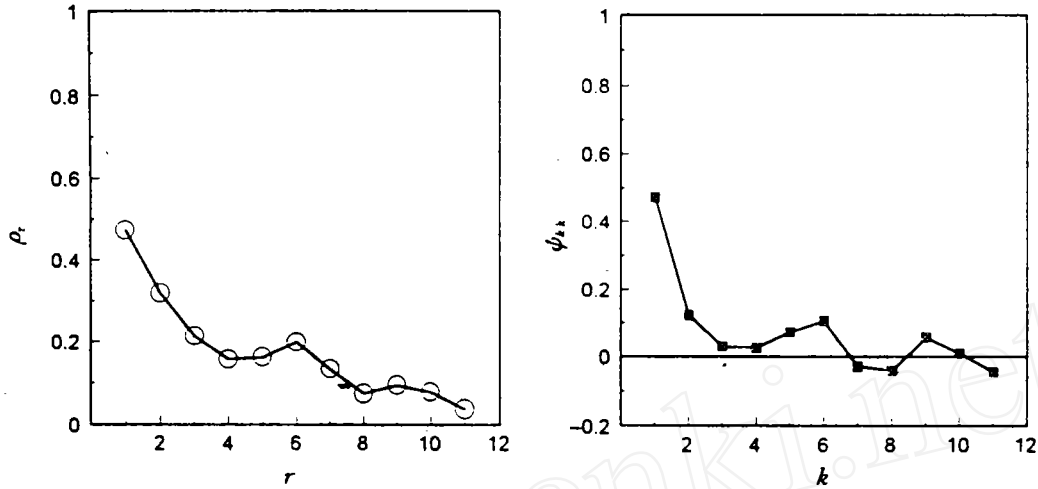


图2 $\{ET_i\}$ 样本的自相关函数(左)和偏相关函数(右)的估计
 Fig. 2 Estimated auto-correlation function (left) and partial auto-correlation function (right) of the $\{ET_i\}$ series

2.2 模型的参数估计(精估计) 模型参数的精估计可采用最小二乘法,它是最常用的一种参数估计方法,这个方法提出的最佳标准是实际值和估计值之差的平方和为最小。用最小二乘法解决各种优化问题时无需假定观察数据的统计性质。根据(1)式,可把 ARMA(p, q)模型表示为:

$$\epsilon_i = ET_i - \varphi_1 ET_{i-1} - \dots - \varphi_p ET_{i-p} + \theta_1 \epsilon_{i-1} + \dots + \theta_q \epsilon_{i-q} \quad (6)$$

利用逆转关系可以将 $\epsilon_{i-j} (j=1, \dots, q)$ 化成 $ET_{i-j}, ET_{i-j-1}, \dots$ 的线性组合,由于逆函数本身是 $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$ 的函数,因此最终得到的方程式对于参数 (φ, θ) 来说必然是非线性的。在一般情况下,(6)式可表示为:

$$\epsilon_i = ET_i - f_i(\varphi, \theta, ET_i) \quad (7)$$

其中 $i=1, 2, \dots, N$ 或 $i=p+1, p+2, \dots, N$, f_i 是对于参数的非线性函数关系。若根据 $N-p$ 个实际观察值可将(7)式表示为:

$$Y = F(ET, \beta) + e \quad (8)$$

其中: $Y = [ET_{p+1}, ET_{p+2}, \dots, ET_N]^T$

$$F = [f_{p+1}, f_{p+2}, \dots, f_N]^T$$

$$e = [\epsilon_{p+1}, \epsilon_{p+2}, \dots, \epsilon_N]^T$$

ET 表示观察值 ET_1, \dots, ET_N , β 表示参数 $\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q$ 。

残差平方和为:

$$Q = [Y - F(ET, \beta)]^T [Y - F(ET, \beta)] \quad (9)$$

由于 F 不是 β 的线性函数,所以根据使 Q 达极小来求得 β 。根据最小二乘原理求得的 AR(1), AR(2), MA(1), MA(2) 及 ARMA(1,1) 模型的参数及残差方差估计 σ_e^2 结果见表 2。从表 2 中得出, ARMA(1,1) 模型的残差方差为最小,但 AR(1) 及 AR(2) 模型的残差方差与 ARMA(1,1) 的残差方差的差异不很显著。因此,需用 AIC 准则选择模型的最佳阶数。

表 2 $\{ET_t\}$ 时间序列的参数估计
Table 2 Estimated model parameters for $\{ET_t\}$ series

参 数 Parameter	ARMA(p, q)模型阶数 / ARMA(p, q) model order				
	AR(1)	AR(2)	MA(1)	MA(2)	ARMA(1, 1)
φ_1	0.471 9	0.427 1			0.670 2
φ_2		0.161 6			
θ_1			-0.365 7	-0.417 0	0.262 8
θ_2				-0.219 1	
σ_ε^2	0.781 7	0.774 8	0.839 2	0.797 8	0.771 4
AIC	-87.9	-89.1	-62.0	-78.5	-90.8

ARMA(p, q)模型最佳阶数的确定可采用 AIC 准则,即给出 AIC 准则函数,该函数既考虑用某一模型拟合时对原始数据接近程度,同时又要考虑模型中所含待参数的个数,建模时按照 AIC 准则函数的取值判断模型的优劣,以决定取舍,使 AIC 准则函数达到极小的为最佳模型。AIC 准则函数可用下式表达:

$$AIC(n) = N \log \sigma_\varepsilon^2 + 2n \quad (10)$$

式中 n 为由模型阶数直接决定的估计参数的个数, σ_ε^2 为在取不同阶数情况下得到的残差方差估计值,在 ARMA(p, q)情况下, $n = p + q$, 此时 AIC 准则函数可定义为:

$$AIC(p, q) = N \log \sigma_\varepsilon^2 + 2(p + q) \quad (11)$$

由表 2 可知, ARMA(1, 1)模型下的 AIC 值为最小,所以本文最后选定的模型为 ARMA(1, 1),其表达式为:

$$ET_t = 0.6702 ET_{t-1} + \varepsilon_t - 0.2628 \varepsilon_{t-1} \quad (12)$$

$$\sigma_\varepsilon^2 = 0.7714$$

3 $\{ET_t\}$ 时间序列的预报

在实际应用中,更关心的问题是如何根据确定的模型及其参数,对未来的蒸散量进行预报,也即根据当前与过去的观察序列的样本值,对该序列未来时刻的取值估计,预报的方法一般采用平稳线性最小方差法。

设 $\{X_t\}$ 是零均值平稳序列,符号 $X_t(L)$ 表示用 t 时刻及其之前的全部历史观察数据,对未来 $t+L$ 时刻的 X_{t+L} 所作的 (L 步)平稳线性最小方差预报,则 $X_t(L)$ 可表示为:

$$X_t(L) = C_0 X_t + C_1 X_{t-1} + C_2 X_{t-2} + \dots \quad (13)$$

被选定的系数序列 $\{C_i\}$ 应使预报误差

$$e_t(L) = X_{t+L} - X_t(L) \quad (14)$$

的均方差达最小:

$$\text{Min}\{E[e_t^2(L)]\} = \text{Min}\{E[(X_{t+L} - X_t(L))^2]\} \quad (15)$$

由 ARMA(p, q)模型白噪声 $\{\varepsilon_t\}$ 的定义可知 $E[\varepsilon_t \varepsilon_{t+r}] = 0$ ($r \neq 0$), 即 ε_t 是相互无关或正交的,因此,可将 $X_t(L)$ 表示或 ε_{t-j} ($j=0, 1, 2, \dots$) 的线性组合:

$$\begin{aligned} X_t(L) &= d_0 \varepsilon_t + d_1 \varepsilon_{t-1} + \dots \\ &= \sum_{j=0}^{\infty} d_j \varepsilon_{t-j} \end{aligned} \quad (16)$$

实现最小线性方差预报的解可表示为:

$$d_j = G_{j+L} \quad (j=0, 1, 2, \dots) \quad (17)$$

因此, L 步线性最小方差预报值为:

$$X_i(L) = \sum_{j=0}^{\infty} G_{L+j} \epsilon_{i-j} \quad (18)$$

上式即是计算 L 步预报值的格林函数(G)方法。在有了 ARMA 模型后, 格林函数可根据模型参数递推计算, 而 ϵ_t 则按(6)式递推计算。上式包括无穷项求和, 实际应用时只需用充分多项的有穷和近似。本文利用建立的 ARMA(1, 1) 模型(见(12)式), 根据 1989 年 1 月 1 日至 3 月 10 日参考作物蒸散量(ET_p)的观测数据(即用 Penman-Monteith 公式计算值), 采用一步预报法预报了 1989 年 1 月 2 日至 3 月 11 日逐日 ET_p 的变化, 其预报结果见图 3。预报误差的方差也即残差方差 σ_ϵ^2 的估计值, 95% 的置信度范围为 $ET_i(1) \pm 1.96\sigma_\epsilon$ 。

4 结论与讨论

①应用自回归滑动平均 ARMA(p, q) 模型估计蒸散量主要考虑了气象因素的随机变化特征, 这种估计方法的主要优点在于对蒸散量的随机变化进行量化, 它考虑了几个变量的滞后值, 也即 t 时刻的变量与 $t-1$ 时刻及其以前的变量有关, 且保证了误差项的独立性, 也即 ϵ_t 与 ϵ_{t-1} 相互独立。这是一般确定性时间序列模型无法做到的。根据建立的 ARMA(p, q) 模型可用于对未来时刻的蒸散量的可能取值进行预报, 并给出预报结果的精度分析。

②本文对蒸散量随机序列 ARMA(p, q) 的低阶模型进行了初步估计和精估计, 并得出了可用 ARMA(1, 1) 随机模型描述 $\{ET_t\}$ 序列的变化特征, 然而, $\{ET_t\}$ 时间序列也有可能为更高阶的 ARMA(p, q) 模型。Box-Jenkins 认为实际问题中 ARMA 模的阶数(p, q) 一般都是比较低的^[6]。因此本文仅对 AR(1), AR(2), MA(1), MA(2) 及 ARMA(1, 1) 进行了估计。

③本文虽然只给出随机序列 $\{ET_t\}$ 的一步预报结果, 在实际应用时, 利用已建立的 ARMA 模型, 还可以用 Kalman 滤波法更精确地预测未来时刻的蒸散量^[4], 也可把它作为土壤水分随机变化的状态空间方程的一部分, 用于预测土壤水分动态变化^[3]。因此, 建立的蒸散量变化的随机模型对制定实时灌水决策和农田水分管理具有重要的参考价值。

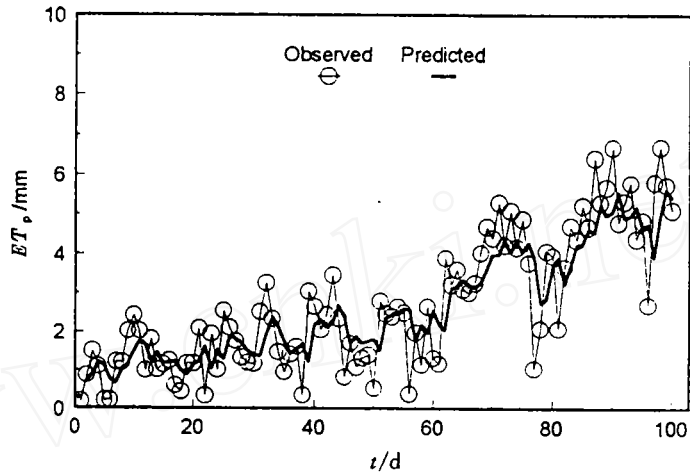


图3 蒸散量观测值和预报值(一步预报法)比较,
(北京, 1989年1月1日~3月10日)

Fig. 3 Comparison between observed and predicted daily ET_p ,
(by one-step forecasting), (Beijing, 1989-01-01~03-10)

参 考 文 献

- 1 杨诗秀, 雷志栋. 均质土壤一维非饱和流动通用程序. 土壤学报, 1985, 22(1): 24~34
- 2 龚元石. 冬小麦和夏玉米农田土壤分层水分平衡模型. 北京农业大学学报, 1995, 21(1): 61~67
- 3 Aboitiz M, Labadie J W, Heermann D F. Stochastic soil moisture estimation and forecasting for irrigated fields. Water Resources Research, 1986, 22(2): 180~190
- 4 Or D, Hanks R J. Irrigation scheduling considering soil variability and climatic uncertainty: Simulation and field studies. In: Russo G, Dagan G eds. Water Flow and Solute Transport in Soils. Advanced Series in Agricultural Sciences, Volume 20. Springer-Verlag, 1993, 262~282
- 5 Allen R G, Jensen M E, Wright J L et al. Operational estimates of reference evapotranspiration. Agron J, 1989, 81: 650~662
- 6 杨位钦, 顾岚编著. 时间序列分析与动态数据建模. 北京理工大学出版社, 1988, 98~106

A Stochastic Model for Evapotranspiration Variability

Gong Yuanshi Li Baoguo

(Dept. of Soil and Water Sciences, CAU, Beijing 100094)

Abstract: Based on stochastic variations of reference crop evapotranspiration (ET_P) time series, a methodology has been developed with a generalized autoregressive moving average (ARMA(p, q)) model for the daily ET_P in this paper. A moment estimation method is applied for the model primary selection. The ARMA(p, q) model parameters are estimated accurately with the method of the least squares. The optimal model order is selected by AIC (Akaika Information Criterion) value. The results showed that the variation of the daily ET_P is represented with a ARMA (1,1) model. Using the developed model, the daily ET_P is predicted for the study site.

Key words: reference crop evapotranspiration; stochastic model