

预报土壤盐分动态的多层递阶时序法^{*}

蔡树英 杨金忠 张瑜芳

(武汉水利电力大学, 湖北武汉 430072)

摘要 文中将多层递阶时间序列方法用于土壤盐分动态的预报,方法简单,计算工作量小,预报精度高,而且避免了利用确定性方法预报时所遇到的困难。实例分析表明,本方法通过对于时变参数的追踪和预报,加上模型本身可修复的特点,使对土壤盐分动态的预测精度得以提高。

关键词 土壤盐分,时间序列,时变参数,多层递阶预报

对于许多大型灌区,引水灌溉为农业生产提供了有力的保障。然而,不合理的灌溉却会给这些地区带来新的灾害,这就是土壤次生盐碱化问题。土壤次生盐碱化的发生发展,取决于土壤含盐量的大小及其随时间的变化,因而,对土壤盐分动态作出及时准确的预报,是控制土壤次生盐碱化产生的必要前提。

土壤盐分动态预报一直是一个较难解决的问题。传统的区域土壤盐分动态预报是从描述盐分在土壤剖面中运移的机理性数学模型出发进行的,即用确定性模型方法研究^[1,2]。确定性模型是将原来复杂的系统进行合理简化后而得出的,尽管如此,还是会出现各种各样计算上的困难。例如,该模型中所用的参数就很难得到。况且,由于土壤结构本身的空间变异性 and 影响土壤盐分运动的各因素,如降雨、蒸发和地下水位等等的随机特性,使得土壤水盐系统为一复杂的随机系统,从而需要非确定性方法来研究这一问题。

时间序列方法就是一种非确定性方法,它通过对观测到的有序随机数据进行分析处理,建立起一种带有参数的数学模型,使其能反映出时间序列本身的物理过程。常用的时间序列模型预报,如AR模型,ARMA模型等的预报,所采用的步骤为:首先通过对测量数据的分析选定模型类型,然后用适用性检验方法及参数识别方法在选定模型的可能阶数范围内确定模型参数。这样确定出的模型类型是时间序列的平均最优模型,而模型参数也是在此最优模型条件下的一组平均最优参数。但是对于实际问题,一般并不知道系统在未来时刻的模型类型(数学模型)是否变化,即使在数学模型不变的情况下,也不知道模型参数是否为常数。因而,寻求一种针对时变参数系统的模型将是十分必要的,多层递阶时序模型正是这样一种模型。

在土壤水盐这个复杂的随机系统中,土壤盐分是受包括土壤水分、降雨、灌溉、蒸发及土壤本身含盐量等各种因素综合影响的结果,其复杂性可想而知。由于影响土壤盐分的诸多因素都是随时间变化的随机量,则土壤水盐系统必定是时变的动态系统。在这个系统中,定期观测得到的野外剖面土壤盐分序列已客观地记载了土壤盐分在各种自然因素

^{*}国家自然科学基金资助项目(批准号:59639240)和211工程青年基金资助项目。

和人为因素作用下变化的历史过程,在其它输入项不易得到的情况下,可以直接利用序列本身建模,由以往的历史规律推测未来,达到预测的目的。本文通过对某大型灌区内各监测点土重百分比含盐量的观测序列进行分析处理,发现其具有复杂的时变特征,故进行了用多层递阶时序模型进行土壤盐分预报的尝试,获得了比较满意的结果。

1 土壤盐分动态的多层递阶数学模型和参数识别

多层递阶时序预报方法的基本思想是把动态系统的状态预报问题分为两部分:即系统时变参数的预报和系统状态的预报^[3],这种方法对原时间序列模型中的参数经过适当处理,又形成一个新的序列,相对于原序列而言,称这些参数形成的时序为第二层时序,同理还可以形成第三层、第四层时序,故称为多层递阶预报方法^[4]。

在土壤水盐这个系统中,我们的目的是将土壤盐分(以土重百分比含盐量表示)作为系统的运转结果,可视为土壤水盐系统的动态输出,而蒸发、降雨和灌溉水的含盐量等因素可视为该系统的动态输入。

将土壤水盐系统表示为单输入、单输出的线性时变系统时,有:

$$S(k) = \varphi^T(k) \cdot \theta(k) \quad k = 1, 2, \dots, N \quad (1)$$

其中, $S(k)$ 是系统的输出; N 为样本容量; $\varphi(k)$ 是观测到的系统输出和系统输入 $\{u(k)\}_1^N$ 所构成的向量:

$$\varphi(k) = \{S(k-1), S(k-2), \dots, S(k-N_a), u(k), u(k-1), \dots, u(k-N_b)\}^T \quad (2)$$

式中, N_a, N_b 为模型输出、输入部分的阶数; $\theta(k)$ 是未知的随机时变参数向量,其估计值 $\{\hat{\theta}(k)\}_1^N$ 可由参数追踪公式^[3,4]求得:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{\delta}{\|\varphi(k)\|^2} \varphi(k) \{S(k) - \varphi^T(k) \hat{\theta}(k-1)\} \quad (3)$$

其中, $k = N_{ab}, N_{ab+1}, \dots, N$; $N_{ab} = \max(N_a, N_b) + 1$;

$$\hat{\theta}(k) = \{\hat{\theta}_1(k), \hat{\theta}_2(k), \dots, \hat{\theta}_{N_a+N_b+1}(k)\}^T \quad (4)$$

$0 < \delta \leq 1$; $\|\varphi(k)\|^2 = \varphi(k)^T \varphi(k)$, 当 $\delta = 1$ 时, 对任意参数初值 $\hat{\theta}(N_{ab-1})$ 所得到的参数序列 $\{\hat{\theta}(k)\}$ 均满足:

$$S(k) = \varphi(k)^T \hat{\theta}(k) \quad (5)$$

也即所求的参数不是唯一的。为了解决这种参数识别的非唯一性,本文采用循环方法选取参数初值^[5,6],首先给定参数初值和序列 $S(k)$, ($k = 1, \dots, N$), 由(3)式追踪得到参数序列 $\hat{\theta}(k)$, ($k = N_{ab}, N_{ab+1}, \dots, N$), 然后由序列 $S(k)$, ($k = N, N-1, \dots, 1$) 和 $\hat{\theta}(N)$ 求得参数序列 $\hat{\theta}^*(k)$, ($k = N, N-1, \dots, N_{ab}$)。设 ε_0 是预先给定的很小的数, 当所有的 k 满足 $\max|\hat{\theta}(k) - \hat{\theta}^*(k)| < \varepsilon_0$ 时, 取向前递推求得的参数作为最优参数的估计值。

2 预报模型和实例

2.1 时变参数和土壤盐分的预报模型

多层递阶预报分为对时变参数的预报和对系统输出的预报两部分。由参数追踪求得

了最优参数估计值 $\{\hat{\theta}(k)\}_1^N$ 以后, 首先检验 $\{\hat{\theta}(k)\}_1^N$ 的平稳性。如果平稳, 则需要针对参数 $\{\hat{\theta}(k)\}_1^N$ 建立相应的模型。由(4)式可以看出, $\{\hat{\theta}(k)\}_1^N$ 是 $N_a + N_b + 1$ 维时间序列, 为简单起见, 按各分量进行处理, 由于对土壤盐分序列所识别出的参数序列具有明显的周期特征, 故对各参数分量 $\{\hat{\theta}_i(k)\}$ 可建立下述组合模型:

$$\hat{\theta}_i(k) = Y_i(k) + P_i(k) + X_i(k) \quad i = 1, 2, \dots, N_a + N_b + 1 \quad (6)$$

$$Y_i(k) = \sum_{L=0}^{s^{(i)}} \alpha_{iL} k^L \quad (7)$$

$$P_i(k) = a_{i0} + \sum_{L=1}^{r^{(i)}} \left[a_{iL} \cos \frac{2\pi k}{T_L^{(i)}} + b_{iL} \sin \frac{2\pi k}{T_L^{(i)}} \right] \quad (8)$$

$$X_{i,k} + \varphi_{i,1} X_{i,k-1} + \varphi_{i,2} X_{i,k-2} + \dots + \varphi_{i,m} X_{i,k-m} = \varepsilon_i \quad (9)$$

式中, $Y_i(k)$ 为幂函数趋势, $P_i(k)$ 为周期项, $T_L^{(i)}$ 为周期, $s^{(i)}, r^{(i)}, \alpha_{iL}, a_{i0}, \alpha_{iL}, b_{iL}$, 为待定参数, $X_i(k)$ 为参数序列减去幂函数趋势和周期趋势后的残差序列。

幂函数趋势的确定采用逐步多项式回归方法; 周期分量 $P_i(k)$ 按以下步骤确定: 首先对 $\{\hat{\theta}_i(k) - Y_i(k)\}$ 序列用方差分析方法或周期图方法在一定检验准则条件下识别其隐含周期 $T_L^{(i)}, L = 1, 2, \dots, r^{(i)}$ 。将 $T_L^{(i)}$ 代入(8)式, 得到线性矛盾方程, 再用最小二乘法确定系数 a_{i0}, a_{iL}, b_{iL} 。残差序列用 AR(m) 模型式(9)描述, 并用最小二乘法识别模型参数, 用 FPE 准则进行模型定阶^[7]。

对时变参数 $\{\hat{\theta}(k)\}_1^N$ 的各个分量建模后, 分别用(7), (8), (9)式对时变参数的各项进行 L 步预报, 代入(6)式即得到时变参数的 L 步预测值 $\{\hat{\theta}^*(k)\}_{N+1}^{N+L}$ 。

在已知系统输出序列 $\{S(k)\}_1^N$, 系统输入序列 $\{u(k)\}_1^{N+L}$, 以及参数序列的追踪值 $\{\hat{\theta}(k)\}_1^N$ 和参数的向前 L 步预报值 $\{\hat{\theta}^*(k)\}_{N+1}^{N+L}$ 的情况下, 就可以通过下面式子得到系统输出序列的向前 h 步预报值:

$$\hat{S}(N+h) = \varphi^*{}^T(N+h) \hat{\theta}^*(N+h) \quad (10)$$

其中,

$$\varphi^*(N+h) = \{\hat{S}_{N+h-1}, \hat{S}_{N+h-2}, \dots, \hat{S}_{N+1}, S_N, S_{N-1}, \dots, S_{N+h-N}, u_{N+h}, u_{N+h-1}, \dots, u_{N+h-N}\}$$

2.2 土壤盐分动态实例预报

实际研究区域是一大型引黄灌区, 位于黄河中下游, 引黄灌溉使该区粮棉油产量大幅度提高。但是历史上曾经由于不适当的灌溉造成该区大面积土壤次生盐碱化发生发展, 加上个别地区存在部分老盐碱地, 虽然经过不断治理, 但土壤次生盐碱化问题一直是该区农业生产的一大威胁。为了有效治理土壤盐碱化和控制次生盐碱化的发生和发展, 灌区从1987年开始设立了土壤水盐监测预报区, 均匀布设了部分盐分观测点, 按季度取土观测, 用于对该区土壤盐分的变化进行监测预报。分析每个测点的实测剖面盐分资料, 考虑植物根系生长范围, 选取各个测点中地面以下0—20cm深度的平均土壤百分比含盐量序列, 建立多层递阶时间序列模型。由于剖面盐分序列本身已经反映了各种因素的综合影响, 加之详细获取系统输入因素的困难, 预报中未考虑系统的输入, 即取 $N_b = -1$, 只用土壤盐分序列本身建模和预报。

文中对区内11个盐分观测点0—20cm深度内平均土壤百分比含盐量序列进行建模

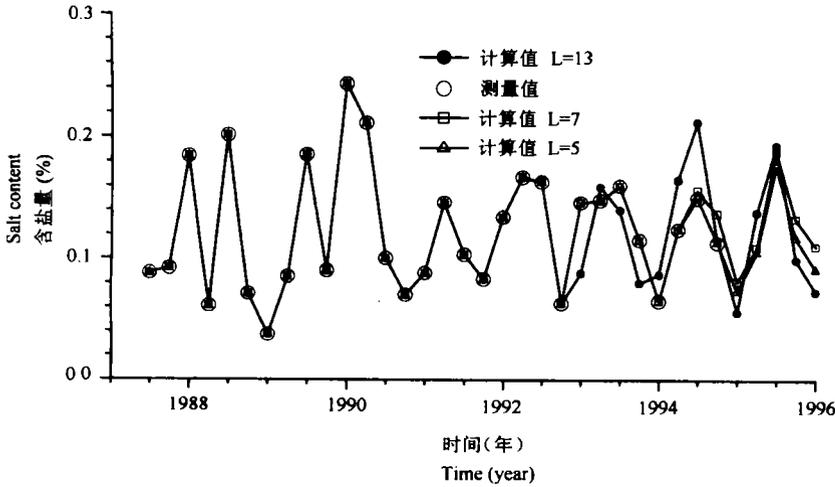


图1 024号点0—20cm深度土壤盐分的实测和计算结果

Fig.1 Measured and calculated soil salinity at the No.024 observation site (0—20cm)

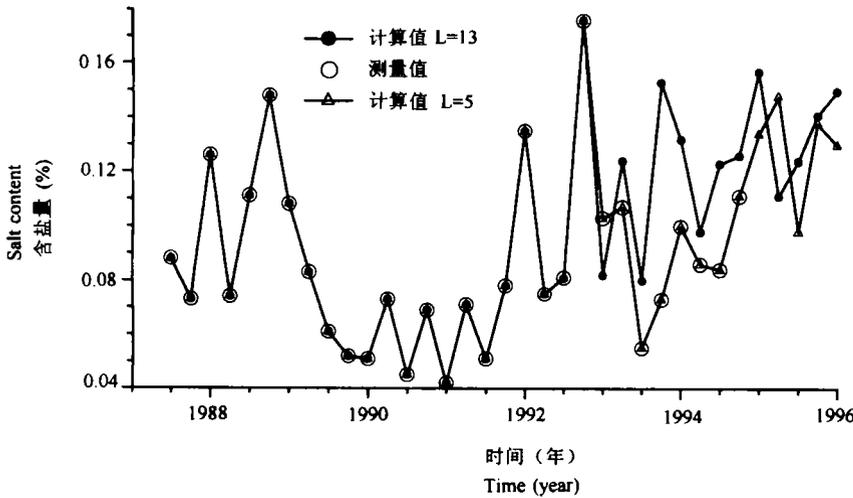


图2 027号点0—20cm深度土壤盐分的实测和计算结果

Fig.2 Measured and calculated soil salinity at the No.027 observation site (0—20cm)

和预报, 由于建模时考虑了序列的时变特性, 通过对时变参数的追踪和预报, 使土壤盐分的预测精度进一步提高。区内 024 号点、027 号点和 703 号点的建模、预报, 并与实测点比较的结果见图 1、图 2 和图 3。在各点的预报中, 利用 1987 年 6 月—1992 年 9 月的实测百分比含盐量序列建模, 从 1993 年初开始进行了 $L = 13$ 步的预报, 其中利用 8 步的预报结果与实测结果进行了比较, 结果看出, 二者基本上是一致的。这说明利用多层递阶时序模型进行剖面土壤盐分的预测预报是完全可行的。

在实际计算中还发现, 时间序列模型具有自身可修复性的特点。随着时间的推移, 新的数据不断补充进来, 模型便可以根据新获得的信息重新建模, 这样修复和改进的模型更

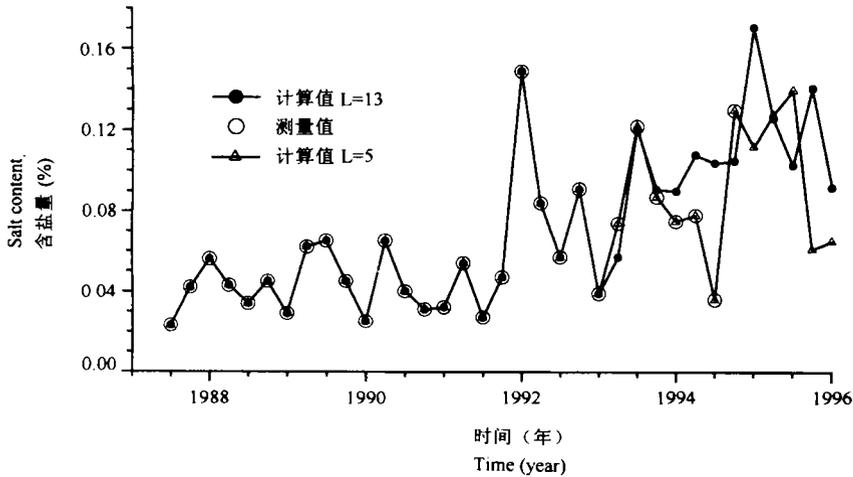


图3 703号点0—20cm深度土壤盐分的实测和计算结果

Fig.3 Measured and calculated soil salinity at the No.703 observation site (0—20cm)

符合于所研究对象的发展趋势,从而使对未来的预测精度得以提高。图1给出了三条建模和预测曲线,其中预测步数 $L = 13$ 步的曲线是从1993年开始预报,可以看出,此曲线对1994年预报结果的偏差明显大于另一条预测步数 $L = 7$ 步的曲线,这是因为后者又将6个新数据点补充给模型,使模型得以修复和改进的结果。而第三条曲线又增加了2个新的数据点,它对后面5步预测的可信度显然应该高于前两条曲线。

同样,图2和图3也给出了一条用于模拟比较的曲线,和将新数据补充进去的预测曲线。

3 结 语

实例预测表明,多层递阶时序方法对于土壤盐分割面的预测是一种行之有效的办法,它避免了利用确定性模型所遇到的困难,所用资料较少且容易得到,模型充分考虑了序列的时变特性,并具有可修复的特点,使预测精度得以提高,而且模型操作简单易行,是一种值得推广的好方法。

参 考 文 献

1. 石元春,李保国,李韵珠,陆锦文. 区域水盐运动预测预报. 河北科学技术出版社,1991.
2. 张蔚榛,杨金忠,张瑜芳,蔡树英. 一种简单的区域水盐动态预测预报方法. 土壤肥料,1992,(5):11—14
3. 韩志刚编著. 多层递阶方法及其应用. 科学出版社,1989. 104,136
4. 蔡季冰编著. 系统辨识. 北京理工大学出版社,1989. 365—366
5. 郑宇宏,韩光文. 多层递阶预报参数估值初值选择的循环方法. 控制理论与应用,1989,6(4):89—95
6. 郑宇宏. 多层递阶预报参数估值初值选择准则及选择方法分析. 控制与决策,1991,6(6):461—464
7. 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. 112—116,141—148

A MULTI-LAYER HIERARCHICAL TIME SERIES MODEL FOR THE PREDICTION OF SOIL SALINITY

Cai Shu-ying Yang Jin-zhong Zhang Yu-fang

(Wuhan University of Hydraulic and Electric Engineering, Wuhan, Hubei 430072)

Summary

A multi-layer hierarchical time series model was used for the prediction of soil salinity. The method was simple and the prediction precision was improved by forecasting and tracking the parameters of the model. The comparison between the predicted and measured results in a large irrigation area showed the good reliability of the method.

Key words Soil salinity, Time series, Time-dependent parameters, Multi-layer hierarchical forecasting