

时间序列分析在土壤墒情预测中的应用研究^{*}

张和喜, 杨 静, 方小宇, 房 军, 冯 诚

(贵州省水利科学研究院, 贵阳 550002)

摘 要: 土壤水分的动态模拟是土壤墒情监测及预报的重要内容, 采用时间序列的线形模型进行土壤水分时间序列的拟合与趋势预测。在建立模型后, 用实测数据与模型的预测数据相比较说明时间序列线形模型能较好地拟合与预测土壤墒情的变化趋势, 可为干旱研究及其治理提供依据。

关键词: 时间序列; 土壤墒情; 预测预报

中图分类号: S152.7

文献标识码: A

文章编号: 1005-3409(2008)04-0082-03

Application of Time Series Analysis in Soil Moisture Forecast

ZHANG He-xi, YANG Jing, FANG Xiao-yu, FANG Jun, FENG Cheng

(Guizhou Academy of Hydraulic Sciences, Guiyang 550002, China)

Abstract: Soil moisture dynamics simulation is an important item of soil moisture monitoring and forecasting. In this paper, time series analysis linear model is used to forecast the change trend of soil moisture. By comparing the measured value with the estimated value, it is showed that the time series analysis linear model can preferably fit soil moisture time series, and provides a foundation for the drought research.

Key words: time series analysis; soil moisture content; forecast

旱灾是我国主要自然灾害之一, 与其他自然灾害相比, 具有发生频繁、涉及面广、历时长等特点, 对农业影响最为严重。在影响旱情分析的诸多因素中, 土壤墒情变化趋势的研究具有重要的意义。通过分析和预测土壤含水率的变化, 可以预知作物的受旱情况, 科学指导农业灌溉和抗旱救灾, 为农业估产和旱灾损失评估提供依据^[1]。

时间序列分析主要是采用参数模型对所观测到的有序的随机数据进行分析与处理的一种方法。目前, 时间序列分析的方法主要应用在医学、金融及各种工程应用领域^[2]。在土壤水分动态模拟方面, 在北方气候条件下康绍忠对较厚土层用提取周期趋势的方法作了土壤水分动态的随机模拟研究, 取得了一定的效果^[3]。在三峡库区, 刘洪斌等人采用时间序列的自回归建模方法对紫色土丘陵旱坡地土壤水分动态进行了模拟和预测^[4]。在本文中为了能够预测土壤墒情的变化趋势, 采用了时间序列的线形模型进行了研究, 并用实测数据与模型的预测数据相比较, 结果显示时间序列线形模型能良好地拟合与预测土壤墒情的变化趋势, 有实际的应用价值。

1 时间序列分析的基本原理

时间序列是以时间先后为分类基准的统计序列, 是指按时间顺序排列的一组数据。时间序列分析的目的就是为了

揭示时间序列本身的结构与规律, 认识系统的固有特性, 掌握系统同外界的关系, 推断系统行为的未来情况^[5]。时间序列预测方法的基本思想是: 预测一个现象的未来变化时, 用该现象的过去行为来预测未来。

1.1 时间序列的 3 种线性模型简介

1.1.1 自回归模型 $AR(p)$

如果随机过程 $X_t(t=0, \pm 1, \dots)$ 的观测值, 与其自身前一个或前 n 个时刻的观测值 X_{t-1}, X_{t-2}, \dots 有关, 或有依赖性, 此时与其前一个时期或前几个时期的观测值之间的回归模型为

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \varepsilon_t \quad (1)$$

式(1)为 P 阶自回归模型, 记为 $AR(p)$, 其中实数 $\varphi_i (i=1, 2, \dots, p)$ 称为自回归系数。

1.1.2 滑动平均模型 $MA(q)$

若 $\{X_t\}$ 为零均值的平稳时间序列

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (2)$$

式(2)为 q 阶滑动平均模型, 记为 $MA(q)$, 其中 $\theta_i (i=1, 2, \dots, q)$ 称为滑动平均系数, 满足 $MA(q)$ 模型的随机序列称为 $MA(q)$ 序列。

1.1.3 自回归滑动平均混和模型 $ARMA(p, q)$

在实际问题中, 更一般的随机序列 $X_t(t=0, \pm 1, \dots)$ 可用式(3)来描述:

^{*} 收稿日期: 2007-05-15

基金项目: 贵州省水利厅主要农作物灌溉专项资金项目

作者简介: 张和喜(1980-), 男, 湖北荆州人, 在读博士, 研究方向: 节水灌溉原理与技术。E-mail: hexi0926@126.com

$$X_t - \alpha_1 X_{t-1} - \alpha_2 X_{t-2} - \dots - \alpha_p X_{t-p} = \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_q \epsilon_{t-q} \quad (3)$$

其中 $\alpha_i (i=1,2,\dots,p)$, $\beta_j (j=1,2,\dots,q)$ 为参数,称上述方程为 P 阶自回归 q 阶滑动平均混和模型,记为 $ARMA(p,q)$ 。它的解 X_t 称为 p 阶自回归 q 阶滑动平均序列^[6]。

2 数据的预分析

首先对原始数据进行数理统计分析。用 2006 年 4 月 26 日至 2006 年 9 月 26 日共获得的 125 组数据进行了初步的统计分析。统计发现土壤水分变异系数的变化有很明显的规律。20,40 cm 深处土壤水分的变异系数接近 30%,属于比较剧烈的变异,40,60,80,100 cm 深处的土壤水分的变异系数在 10%~20%之间,属于中等程度的变异,而在 120,140 cm 处变异系数小于 5%,属于弱变异。这说明越浅层的土壤水分值变化越快,不稳定,而深层土壤水分值却变化缓慢,比较稳定。这样在做时间序列的分析时将考虑用不同的时间序列模型来拟合不同埋设深度的土壤水分变化。

3 时间序列模型结果与分析

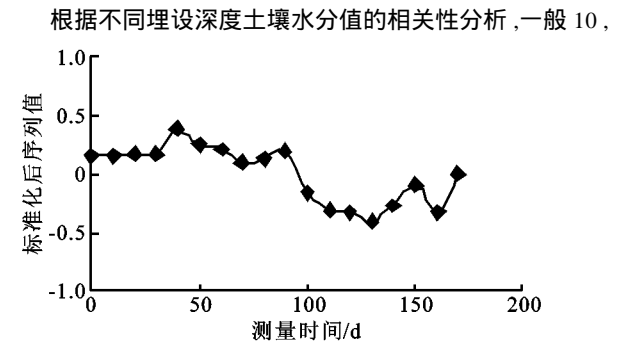


图 1 标准化后 60 cm 埋深处土壤水分值

由图 2 可知,一阶差分后序列的均值比较接近于零,序列没有明显的时间变化趋势。可以认为一阶差分后的土壤水分序列为均值是零,平稳的数据序列,符合时间序列的建模要求。

3.2 模型阶数确定

由相关函数的性质及计算误差精度确定预报方程阶数,实际应用时预报阶数可取:

$$m = \min\{N - 1, \lceil N/4 \rceil\} \quad (6)$$

3.3 最优线形预报方程的建立

$$\begin{bmatrix} B(0) & B(1) & B(2) & B(3) & \dots & B(m-1) \\ B(1) & B(0) & B(1) & B(2) & \dots & B(m-2) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ B(m-1) & B(m-2) & B(m-3) & B(m-4) & \dots & B(0) \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ \dots \\ m \end{bmatrix} = \begin{bmatrix} B(1) \\ B(2) \\ \dots \\ B(m) \end{bmatrix} \quad (9)$$

将试验测定数据代入所建时间序列模型中,得
 $B(0) = 0.026344, B(1) = -0.00618, B(2) = 0.000771367,$

$$\begin{bmatrix} 0.026344 & -0.00618 & 0.000771 & 0.005892 \\ -0.00618 & 0.026344 & -0.00618 & 0.000771 \\ 0.000771 & -0.00618 & 0.026344 & -0.00618 \\ -0.005892 & 0.000771 & -0.00618 & 0.026344 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -0.00618 \\ 0.000771 \\ 0.005892 \\ -0.00666 \end{bmatrix} \quad (10)$$

40,60,100 cm 处的土壤水分值具有代表性的意义^[6]。在本文中将以 60 cm 处的土壤水分数据为例进行分析,建立时间序列模型。

3.1 数据的整理

为了满足数据分析的需要,同时提高运算精度,首先应对原始数据进行标准化处理,标准化公式为

$$Y_t = \frac{X_t - u}{v} \quad (4)$$

式中: Y_t ——标准化后的时间序列; X_t ——原始时间序列; u ——时间序列 $\{X_t\}$ 的均值; v ——时间序列 $\{X_t\}$ 的标准差,标准化后的数据见图 1,图中横坐标为测量时间,纵坐标为标准化后 60 cm 埋深处的土壤含水量。

时间序列模型的分析都是建立在序列平稳的条件上的。一个平稳的随机过程有以下要求:均值不随时间变化;自相关系数只与时间间隔有关,而与所处的时间无关。实际上,大多数时间序列都是不平稳的。由图 1 可知,标准化后 60 cm 深处土壤水分值的序列显然也不满足平稳的要求,需要对序列进行差分处理,差分处理的公式如下:

$$DY_t = Y_t - Y_{t-1} \quad (5)$$

式中: DY_t ——一阶差分; Y_{t-1} ——差分后的序列通常表现出没有明显的时间趋势以及剧烈波动等特征。

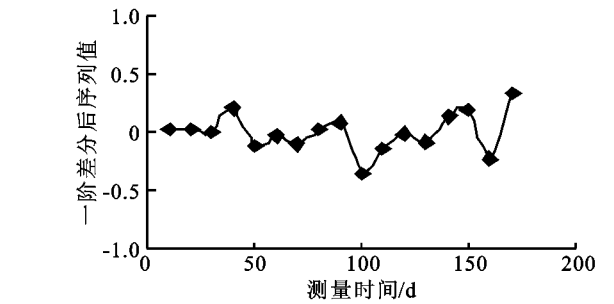


图 2 差分处理后的序列

在最小二乘法原则的基础上,确定线形最优预报方程:

$$Y_{n+1} = \alpha_1 Y_n + \alpha_2 Y_{n-1} + \dots + \alpha_m Y_{n-m+1} \quad (7)$$

其中 α_i 取正整数(外推步长)。

3.4 计算

$$B(k) = \frac{1}{N-K} \sum_{i=1}^{N-K} y_{i+k} y_i \quad (8)$$

计算 $B(0), B(1), \dots, B(m)$ 。

3.5 建立并解 Yule - Walker 方程

由此可得 Yule - Walker 方程的线形方程组

$$\begin{bmatrix} 0.026344 & -0.00618 & 0.000771 & 0.005892 \\ -0.00618 & 0.026344 & -0.00618 & 0.000771 \\ 0.000771 & -0.00618 & 0.026344 & -0.00618 \\ -0.005892 & 0.000771 & -0.00618 & 0.026344 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -0.00618 \\ 0.000771 \\ 0.005892 \\ -0.00666 \end{bmatrix} \quad (10)$$

运用 Matlab 软件对 Yule - Walker 方程进行数值处理解得：

$\gamma_1 = -0.1957, \gamma_2 = -0.0349, \gamma_3 = -0.1993, \gamma_4 = -0.1633$

预测 2006 年 10 月土壤墒情将上述 $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ 的数值代入预报方程 (1) 得自回归预报模型：

$$y_{18} = -0.1957y_{17} + 0.0349y_{16} + 0.1993y_{15} - 0.1633y_{14} = 0.0582 \tag{11}$$

又根据式 (5) $Y_{19} = 0.2724$

再又公式 (4) $X_{19} = 14.47$

以上预测土壤墒情与实际所测土壤含水量相吻合, 预测精度较高。同理根据此预报模型可预测实验地区未来的土壤墒情。应用 SPSS 软件中的自回归模型, 建立实验地区土壤墒情预测模型, 得到预测分析结果如表 1。

表 1 预测分析结果表				%
序号	月份	实测值	预测值	相对误差
1	10 月中旬	16.25	18.23	12.185
2	10 月下旬	15.02	14.19	- 5.526
3	11 月上旬	18.39	18.21	- 0.979
4	11 月中旬	16.86	17.25	2.313
5	11 月下旬	17.59	16.53	- 6.026

从表 (1) 中可以看出, 用时间序列分析建立实验地区土壤墒情预测模型, 预测值十分接近观测值, 说明时间序列分析适合于土壤墒情的模拟与预测。

4 结 语

时间序列分析模型实际上是一种自回归模型, 即未来数

据与历史数据之间存在自相关关系, 建立反映这种自相关关系的模型就可以对系统的未来情况进行预测。为了提高土壤墒情预报结果的准确性, 在实际应用中可采取多种预测方法, 相互验证。但其不能反映未来长时间的土壤墒情, 因而只适合对未来趋势做短期预测。但它对资料的要求比较单一, 只需要变量本身的历史数据, 因而在实际情况中有着广泛的适用性。本文利用时间序列线形方法对实验区土壤墒情建立预测模型, 从预测结果可以看出, 所建模型具有较好的适应性和预报精度且拟合效果较好, 说明这种预测方法有一定的实用性, 并对今后开展其他地区土壤墒情预报提供科学依据。

参考文献：

[1] 杨绍辉, 王一鸣, 等. ARIMA 模型预测土壤墒情研究[J]. 干旱地区农业研究, 2006, 24(2) : 114-118.
[2] 杨叔子, 吴雅. 时间序列分析的工程应用[M]. 武汉: 华中理工大学出版社, 1991 : 1-395.
[3] 康绍忠. 土壤水分动态的随机模拟研究[J]. 土壤学报, 1990, 27(1) : 17-24.
[4] 刘洪斌, 武伟, 魏朝富, 等. AR 模型在土壤水分动态模拟中的应用[J]. 山地学报, 2004, 22(1) : 121-125.
[5] 迟道才, 张宁宁, 等. 时间序列分析在辽宁朝阳地区干旱灾变中的应用[J]. 沈阳农业大学学报, 2006, 37(4) : 627-631.
[6] 孙凯. 墒情(旱情) 监测与预测预报方法研究[D]. 北京: 中国农业大学, 2004.

(上接第 81 页)

参考文献：

[1] 王伯荪, 王昌伟, 彭少麟. 生物多样性刍议[J]. 中山大学学报, 2005, 44(6) : 68-70.
[2] 吴征镒. 中国植被[M]. 北京: 科学出版社, 1980 : 956-960.
[3] 刘瑛心. 我国荒漠植物区系形成的探讨[J]. 植物分类学报, 1982, 20(2) : 131-142.
[4] 张文辉, 康永祥, 李红. 西北地区生物多样性特点及其研究思路[J]. 生物多样性, 2000, 8(4) : 422-428.
[5] 王涛. 中国沙漠与沙漠化[M]. 石家庄: 河北科学技术出版社, 2003 : 162-263.
[6] 王思博. 新疆啮齿动物志[M]. 乌鲁木齐: 新疆人民出版社, 1998.

[7] 王香亭. 甘肃脊椎动物志[M]. 兰州: 甘肃科学技术出版社, 1991.
[8] 马世威, 马玉明, 姚洪林. 沙漠学[M]. 呼和浩特: 内蒙古人民出版社, 1998.
[9] 陈灵芝. 中国的生物多样性现状及其保护对策[M]. 北京: 科学出版社, 1994.
[10] 李澍卿, 周伟文, 田翠琴. 旱区环境社会学[M]. 石家庄: 河北人民出版社, 2001 : 37-38.
[11] 袁国映, 王琳, 卓丽菲娅. 西北干旱区主要濒危动植物现状及其 21 世纪保护措施展望[J]. 新疆环境保护, 1996, 18(2) : 44-47.
[12] 张维平. 保护生物多样性[M]. 北京: 中国环境科学出版社, 2001 : 15-17.