

基于神经网络的混沌时间序列土壤墒情预测预报

邓建强, 陈效民, 方堃, 杜臻杰

(南京农业大学 资源与环境科学学院, 江苏 南京 210095)

摘要: 土壤墒情预测预报对农业生产、水分循环的研究具有重要的意义。应用混沌理论对具有混沌特性的土壤墒情时间序列进行相空间重构, 利用神经网络对土壤墒情时间混沌序列重构相空间中相点的演化过程进行了学习、训练及预测。结果表明, 该方法所需的参数较少, 简单易行, 即只需要土壤墒情时间序列数据。通过对预测预报值与实测数据进行比较, 证实了该方法相对误差较小, 预测精度高, 有一定的可靠性和实用性。

关键词: 土壤墒情; 混沌; 神经网络; 预测预报

文献标识码: B

文章编号: 1000-288X(2008)06-0082-04

中图分类号: S152.7

Prediction of Chaotic Soil Moisture Time Series Based on Artificial Neural Network

DENG Jian-qiang, CHEN Xiao-min, FANG Kun, DU Zhen-jie

(College of Resources and Environmental Science, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: The prediction of soil moisture is significant to the research on agricultural production and water cycles. Artificial neural network is used to approximate the phase space reconstruction of chaotic soil moisture time series and the future soil moisture was then predicted. Results show that this method is easier to be used in practice because it only needs one parameter - soil moisture time series. The comparison between the predicted value and the measured value indicates that the prediction method has a little relative error and better prediction accuracy. The study also demonstrates the utility and efficiency of the method for predicting soil moisture.

Keywords: soil moisture; chaos; artificial neural network; prediction

土壤墒情是作物生长、植被恢复的关键性限制因子。如何准确、快速预测土壤墒情对农业生产、水分平衡及土壤—植物—大气水分循环的研究具有重要意义。目前主要预测方法有引入随机变量预测法, 土壤水分平衡方程估算法, 土壤水动力学预测法等^[1]。但是以上方法在实际的预测中具有所需参数多, 人力物力投入大, 预测过程复杂, 难以实际运用等缺点。因此, 寻求一种简单有效的土壤水分预测方法具有重要的理论和实际意义。

随着混沌系统和神经网络理论的发展, 这两种理论越来越受到各学科的关注。由于该预测方法具有预测过程简单的特点, 此方法被应用于经济、电力、水文和农业等系统的时间序列预测^[2-6]。土壤墒情受多因素的影响, 具有一定的混沌特征。本文采用基于神经网络的混沌时间序列预测法对研究区土壤墒情

进行预测, 以便为研究区的水分的合理利用和水土保持提供科学依据。

1 预测原理

1.1 相空间重构

根据混沌时间序列去研究复杂系统的基础是相空间重构。其基本理论是: 由于混沌系统产生的轨迹经过一定时期的变化后, 会最终作一种有规则地运动, 产生一种规律的、有形的轨迹。根据 F. Takens 和 R. Mane 的延迟嵌入定理^[7-8], 即选取适当延滞时间和嵌入维 m 重新构建相空间, 其原混沌动力系统的几何特征与重构相空间的几何特征便是等价的, 它们具有相同的拓扑结构。这意味着可把预测问题转化为相空间里的一个短的演化过程来讨论, 从而为混沌时间序列的预测奠定坚实的理论基础^[9]。

收稿日期: 2008-05-19

修回日期: 2008-08-31

资助项目: 国家重点基础研究发展计划(973 计划)(2005CB121103); 中国科学院南京土壤研究所土壤与农业可持续发展国家重点实验室开放课题(0751010015)

作者简介: 邓建强(1984—), 男(汉族), 内蒙古乌兰察布市人, 在读硕士, 主要从事水土资源利用与管理方面的研究。E-mail: djq1dq2@sina.com。

通信作者: 陈效民(1957—), 男(汉族), 江苏省张家港市人, 博士, 教授, 主要从事水土资源利用及管理及模型模拟方面的研究。E-mail: xmrchen@njau.edu.cn。

设单变量的土壤墒情时间序列为 $\{X(t_i)\}$

$$X(t_i) = X(t_0 + n \cdot t) \quad (i = 1, 2, 3, \dots, n) \quad (1)$$

式中: t_0 ——初始时间; t ——采样时间间隔。

选择合适的延滞时间 (τ) 和嵌入维数 (m) 对时间序列进行重构一个 m 维相空间 $\{Y(t_i)\}$

$$Y(t_i) = [X(t_i), X(t_i + \tau), \dots, X(t_i + (m - 1) \cdot \tau)] \quad (i = 1, 2, 3, \dots, m) \quad (2)$$

公式(2)中, 重构相空间 $\{Y(t_i)\}$ 中有 M 个相点 $[M = n - (m - 1) \cdot \tau]$, 每个相点有 n 个分量 $[n = 1 + (m - 1) \cdot \tau]$ 。在此 M 维相空间中, 相点间的连线描述了系统在 M 维相空间的演化轨迹。

1.2 关联维数

混沌现象最主要的特征是空间中存在着奇异吸引子, 而描述吸引子的基本数学量是它的维数。最常见的维数包括: Hausdorff 维数、信息维数、Lyapunov 维数、关联维数等。其中最重要的维数是关联维数。规定凡距离小于给定正数 r 的向量, 称为有关联的向量, 假定一共构造了 N 个向量 y_i , 并定义相关积分函数为

$$C(r) = \frac{1}{N^2} \sum_{i,j=1}^N \chi(r - r_{ij}(x)) \quad (3)$$

式中: $x = r - r_{ij}$; $\chi(x)$ ——Heavisid 单位函数。

$$\chi(x) = \begin{cases} 0 & (x \leq 0) \\ 1 & (x > 0) \end{cases} \quad (4)$$

在公式(3)中, r_{ij} 为所重构相空间中任意两向量的标准欧基里德范数。在 r_{ij} 的最大值和最小值之间调整 r 值的大小, 算出一组关于 $\ln r$ 及 $\ln C(r)$ 。我们把关联维数 D 定义为

$$D(m) = \lim_{r \rightarrow 0} \frac{\ln C(r)}{\ln(r)} \quad (5)$$

1.3 Lyapunov 指数

轨道的收敛率或发散率称为 Lyapunov 指数, 它是研究混沌的一个重要参数, 即是根据轨迹间的距离的无扩散运动特征来判别系统的混沌特性^[10]。但是并不是需要计算出序列的所有 Lyapunov 指数谱, 而只要计算出最大 Lyapunov 指数。本文运用 Wolf 法来计算最大 Lyapunov 指数。即为直接基于轨相线、相平面、相体积等演化来估计 Lyapunov 指数的方法。

设混沌时间序列为 $\{X(t_i)\}$ 如公式(1)所示, 嵌入维数 m 和延滞时间 τ 对时间序列进行重构一个 m 维相空间为 $\{Y(t_i)\}$ 如公式(2)所示。取初始点 $Y(t_0)$, 设与其最近邻点 $Y_0(t_0)$ 的距离为 L_0 , 追踪这两点的演化, 直到 t_1 时刻, 其间距离超过某规定值 ϵ , $L_0 = |Y(t_1) - Y_0(t_1)| > \epsilon$, 保留 $Y(t_1)$, 并在 $Y(t_1)$ 邻近另找一个点 $Y_1(t_1)$, 使得 $L_1 = |Y(t_1) - Y_1(t_1)| < \epsilon$, 并且与之夹角尽可能的小, 继续上述过程, 直至 $Y(t)$ 到达时间序列的终点 N , 这时追踪演化过程

总的迭代次数为 M , 最大 Lyapunov 指数如公式(6)所示。

$$\lambda_1 = \frac{1}{M} \sum_{i=0}^{M-1} \ln \frac{L_{i+1}}{L_i} \quad (6)$$

1.4 BP 神经网络

神经网络具有强大的非线性映射能力, 所以它可以有利地逼近混沌时间序列重构相空间中相点的演化过程。目前比较成熟且应用广泛的网络为 BP(back-propagation) 神经网络。该网络由输入层、隐含层及输出层组成。每层都可以包含若干个神经元, 同一层各神经元之间没有相互连接, 相邻层的神经元之间通过权连接。BP 神经网络是一种误差反向传播的多层前馈式网络, 即将网络输出误差反向传播对网络权重及阈值进行修正, 从而实现网络的映射能力。最基本的网络结构如图 1 所示, 具体算法见文献[11]。

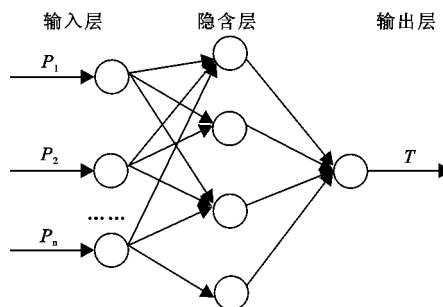


图 1 三层 BP 神经网络结构图

2 建模方法

本研究使用 MATLAB7.5 建立预测模型。建模过程包括: 数据收集及整理, 相空间重构, 最大 Lyapunov 指数计算, 网络设计, 网络训练, 网络预测。

2.1 研究区概况及数据收集

研究区位于湖南省南部的中国农科院祁阳红壤实验站(东经 111°53', 北纬 26°45'), 属于典型的红壤丘陵区。海拔高度 100 ~ 300 m, 年平均气温 18℃, 年降雨量 1 290 mm, 无霜期 295 d, 年日照约 1 613 h。研究区的面积为 216 m², 共设置了 12 个小区, 每个小区为 18 m² (3 m × 6 m, 每个小区中安装两根中子管)。试验从 2007 年 3 月到 2008 年 3 月, 每 7 d 利用中子仪测定一次 0—25 cm 土壤耕作层土壤含水量。由于耕作层土壤墒情变化受外界条件影响大, 且该层次的土壤墒情变化对植被的生长有较强的影响, 所以本研究选取旱田的 5 个小区中变化稳定且相近的耕作层的土壤含水量作为基础数据, 后对异常数据进行剔除, 取其平均值进行模拟。土壤耕作层的基本理化性质如表 1 所示。

2.2 相空间重构及最大 Lyapunov 指数的计算

本文采用自相关函数法来计算延滞时间，即土壤墒情时间序列的自相关函数下降到初始值的 $1 - 1/e$ 时为所确定时间延迟。通过表 2 自相关系数可确定延滞时间 $R = 1$ 。饱和关联维数 D 采用 G-P 算

法来计算,即随着嵌入维 m 的增加其对应关联维数 D 随之变化,当其达到稳定时为所求的饱和关联维数 D 。根据表 3 可知,当相空间维数 m 增加到 5 时, D 趋于稳定其值为 2.60,由此系统的最佳嵌入维 $m = \text{int}(2D) + 1 = 6$ 。

表 1 供试土壤的基本性质

pH 值	有机质含量/ (g · kg ⁻¹)	容重/ (g · cm ⁻³)	总孔隙度/ %	土壤质地/(g · kg ⁻¹)		
				黏粒含量	粉砂粒含量	砂粒含量
5.12	9.02	1.16	56.23	357.31	513.48	129.21

表 2 自相关系数 R 随延滞时间 的变化

自相关系数	1	2	3	4	5	6	7	8
延滞时间 R	0.582 4	0.313 4	0.188 1	0.321 4	0.197 3	0.074 9	- 0.057 0	- 0.024 1

表 3 关联维数 D 随嵌入维数 m 的变化

m	2	3	4	5	6
D	1.459 6	2.183 7	2.499 0	2.603 5	2.533 8

Lyapunov 指数 (λ_1) 是序列混沌特性的显著标志,一个系统是否混沌可以用它的 λ_1 是否大于 0 来判断。本文由 Wolf 法得 $\lambda_1 = 0.151 6$, 因为 $\lambda_1 > 0$, 说明该系统是具有混沌特性。在实际的应用中,系统的最大可预报时间尺度可利用 λ_1^{-1} 进行估算^[12]。根据上述公式可得到该土壤墒情的预测时间尺度约为 2 个月。

2.3 网络设计及训练

理论及实践证明,3 层 BP 网络可以逼近任何有理函数。本研究中对 3 层 BP 网络进行训练,传递

函数各层均采用双曲正切 (Tansig) 传递函数,训练函数采用一步割线算法 (TRAINOSS), 该算法的优点是收敛速度较快。自适应学习函数采用梯度下降学习算法 (LEARN GDM) 对网络进行训练,通过调整隐含层神经元个数来确定网络结构,最终确定隐含层神经元个数为 6 时,网络具有误差收敛速度快,拟合误差小,泛化能力强的特点。所以该 BP 网络模型的拓扑结构为 5 6 1,其中输入及输出层分别为每组相重构空间中的前 5 个向量及最后 1 个向量。综上所述,经过相空间重构得到 43 组重构相量。以前 33 组重构相量作为训练样本对网络进行训练,以后 10 组重构相量作为预测样本,对该方法的可行性及预测精度进行判别。表 4 为经过训练后所得网络权重及阈值。

表 4 BP 神经网络权重(W)及阈值(B)

序号	输出层				输入层			
	W_1	B_1	W_2	B_2	W_1	B_1	W_2	B_2
1	- 2.465 0	1.761 0	2.511 8	- 1.672 4	0.000 3	0.503 5	2.250 3	
2	- 3.573 6	0.675 6	0.177 5	- 0.828 9	0.665 4	2.211 9	- 2.851 6	
3	- 1.289 1	2.029 4	- 1.381 8	- 3.678 5	0.619 6	0.034 6	1.988 0	0.019 7
4	- 0.554 5	0.247 4	0.208 3	2.099 6	0.837 9	- 0.613 1	4.025 6	
5	- 0.236 1	2.028 4	2.497 6	1.578 1	0.899 2	- 2.087 2	- 2.297 7	
6	- 0.402 8	- 0.820 1	3.386 2	- 2.037 1	- 1.232 1	- 1.857 4	- 1.853 0	

注: W_1, B_1 为输入层权重及阈值; W_2, B_2 为输出层权重及阈值。

2.4 网络预测

本文对 2007 年 3 月到 2007 年 12 月数据进行网络训练,对 2008 年 1 月初到 3 月初土壤墒情进行预测预报。图 2 和图 3 分别为土壤墒情预测图及其相应的相对误差图。由图 2—3 可知,基于神经网络的

混沌时间序列土壤墒情预测具有一定的可行性及可靠性。对开始一月初到二月中旬土壤墒情预测的相对误差比较平稳且均小于 3%,由原基础数据相对误差统计(表 5)可知该时间段预测相对误差在允许误差范围之内。但是随着时间尺度的延长预测相对误

差忽高忽低处于不平稳的阶段,因此该预测的最大时间尺度大约一个半月,适合于短期的土壤墒情预测。

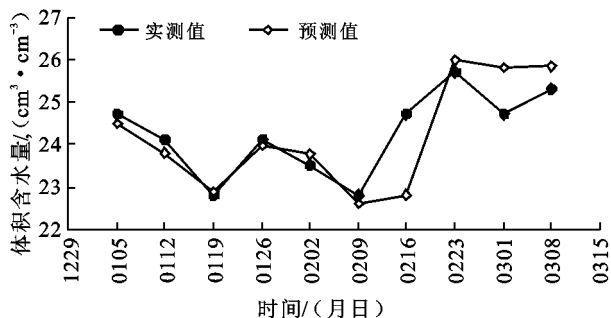


图 2 土壤墒情预测图

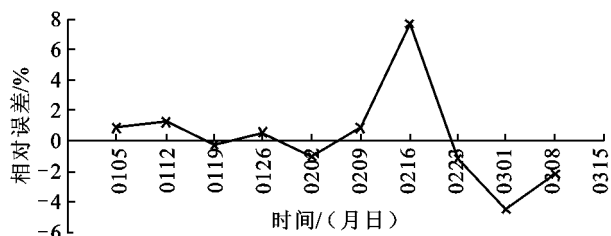


图 3 土壤墒情预测的相对误差图

表 5 原基础数据相对误差统计

相对误差	1 %	2 %	3 %	4 %
累积贡献率	21 %	53 %	91 %	100 %

3 结论

土壤墒情时间序列具有混沌特性,应用 Wolf 算法确定最大 Lyapunov 指数为 0.1516;通过自相关函数法计算得延滞时间 为 1;采用 G—P 算法确定饱和关联维数(D)等于 2.60,最佳嵌入维 m 为 6;根据延滞时间 及最佳嵌入维 m 对相空间进行重构。利用 BP 神经网络对土壤时间混沌序列重构相空间中相点的演化过程进行学习、训练及预测。结果显示,该方法需要参数少,简单易行,即只需要土壤墒情时间序列数据;预测相对误差小于 3%,证明该方法具

有可靠性和可行性;但是由于试验时间序列数据量有限,本预测模型只适用于短期预测,最大预测尺度为一个半月。

[参 考 文 献]

- [1] 申慧娟,严昌荣,戴亚平. 农田土壤水分预测模型的研究进展及应用[J]. 生态科学,2003,22(4):366—376.
- [2] 于国荣,夏自强. 混沌时间序列支持向量机模型及其在径流预测中应用[J]. 水科学进展,2008,19(1):116—122.
- [3] 陈朝霞,周宏,余德贵. 基于混沌吸引子的复杂农业系统预测模型研究[J]. 南京气象学院学报,2004,27(6):862—866.
- [4] 张玉梅,曲仕茹,温凯歌. 基于混沌和 RBF 神经网络的短时交通流量预测[J]. 系统工程,2007,25(11):26—30.
- [5] 张桂英. 基于神经网络的混沌时间序列短期预测[J]. 计算机工程,2002,28(11):197—198.
- [6] 程广平,汪波. 基于神经网络的混沌系统状态预测[J]. 系统仿真学报,2007,19(5):1173—1175.
- [7] Takens F. On the numerical determination of the dimension of an attractor[C]// Rand D, Young L S. Dynamical systems and turbulence. Warwick, 1980. Lecture Notes in Mathematics:Spring-Verlag,1981,898:366—381.
- [8] Mane R. On the dimension of the compact invariant sets of certain nonlinear maps[C]// Rand D, Young L S. Dynamical systems and turbulence. Warwick, 1980. Lecture Notes in Mathematics:Spring-Verlag, 1981, 898:230.
- [9] 王海燕,卢山. 非线性时间序列分析及其应用[M]. 北京:科学出版社,2006:12—28.
- [10] 付强. 数据处理方法及其农业应用[M]. 北京:科学出版社,2006:370—385.
- [11] 张代远. 神经网络新理论与方法[M]. 北京:清华大学出版社,2006:22—26.
- [12] 韩敏. 混沌时间序列预测理论与方法[M]. 北京:中国水利水电出版社,2007:57—61.