

基于改进 BP 算法的地下水动态预测模型

卢文喜 杨忠平 李 平 杨 威

(吉林大学环境与资源学院,吉林 长春 130026)

摘要 运用学习率自适应动量 BP 算法建立了吉林西部地下水埋深人工神经网络模拟预测模型。首先利用自回归分析方法确定网络输入输出样本,而后应用“试错法”确定隐含层节点数,最终建立了 6:10:1 的 ANN 地下水动态模拟预报模型,最后应用 VB 语言依据改进 BP 算法编制计算程序进行模拟计算。通过对模型检验可知该模型模拟和预测精度均较高,完全可应用于地下水位动态预报。2002 年以后的预报结果表明该地区地下水位持续下降,应及时加以控制。

关键词 人工神经网络 改进 BP 算法 地下水动态 动态预报 吉林西部

中图分类号 :TV211.1 文献标识码 :A 文章编号 :1004-693X(2007)03-0005-04

Dynamic prediction model of groundwater level based on improved BP algorithm

LU Wen-xi, YANG Zhong-ping, LI Ping, YANG Wei

(College of Environment and Resources, Jilin University, Changchun 130026, China)

Abstract A groundwater depth prediction model of artificial neural network (ANN) for West Jilin was established based on a self-adapted BP algorithm. First, the input and output samples for the network were determined through autoregression analysis, then the hidden units using the trial-and-error method and an ANN model with a structure of 6:10:1 were determined for the simulation and prediction of dynamics of groundwater; finally, a computer program was made with VB according to the improved BP algorithm. The validations of the model show that the precision of the simulation and prediction is high. This model can be applied to the forecast of groundwater dynamics. The predictions after 2002 indicate a continuing decline of groundwater level in the regions, which should be controlled in time.

Key words artificial neural network (ANN); improved BP algorithm; dynamics of groundwater; dynamic prediction; West Jilin

地下水动态受自然和人为因素的影响,常表现出复杂的非线性特征。建立地下水位动态预测模型是利用了人工神经网络(ANN)的处理非线性模式识别特性。模型的建立主要依赖常规的水文气象观测资料,无需再进行专门的试验来获取参数,节省人力物力,给地下水动态预报分析带来了便利。

ANN 模型自 1943 年兴起以来,在各个领域得到了较好应用,但直到 20 世纪 90 年代才引入到水科学中来,尤其是近几年来,随着计算机技术的发展,已广泛应用于水文分析、预测以及水质评价等领域,取得了较好效果,并发展了多种模型^[1-4],其中 BP

(Back Propagation)网络模型应用最多,但在地下水动态预测中大多还停留在起步应用和试探阶段,远不如在陆地水文、水质评价等方面的应用广泛和深入^[5]。

鉴于此,本文探讨应用一种改进 BP 网络模型——学习率自适应动量 BP 网络模型对吉林西部地下水位动态进行分析,以揭示该地区地下水动态变化规律,提供较为准确的地下水位动态模拟和预报数据,为吉林西部地下水合理开发和科学管理提供依据,同时为 ANN 模型技术在区域地下水动态模拟预测中的推广应用提供借鉴。

1 BP 网络基本原理和算法实现

1.1 基本原理

BP 网络实际是一种误差反向传播的多层前馈型人工神经网络,是用工程技术手段模拟人脑神经网络的结构和功能的一类非线性动力学系统。通常由输入层、若干隐含层和输出层组成。每层都包含若干节点(即神经元),它是 BP 网络的基本处理单元,其从前一层神经元接受信息,经过神经元内部转换后,向下一层神经元发出信息,从而通过神经元之间的相互作用来完成整个网络的信息处理^[6-7]。其网络拓扑结构如图 1 所示。

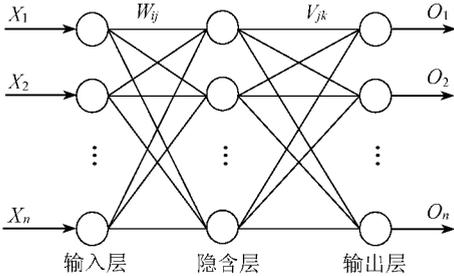


图 1 BP 网络拓扑结构示意图

图 1 为典型的三层 BP 网络,同一层各神经元相互无连接,相邻层的神经元通过权实现全连接。BP 算法实际是前向式多层网络的实际输出与期望输出之间最小均方差的一种迭代梯度算法。其学习过程由正向传播和反向传播组成,正向传播过程中,输入信号 x_i 从输入层经隐层单元逐层处理后传向输出层,每一层神经元的状态只影响下一层神经元的状态,如果在输出层得不到期望值,则转入反向传播,将误差信号沿原连接通路返回,通过修正各层神经元的权值和阈值使网络全局误差信号达到给定的精度标准,从而完成学习过程。这样,当在网络的输入端加入一个新的信号时,就能从其输出端得到相应结果。如图 2 所示。

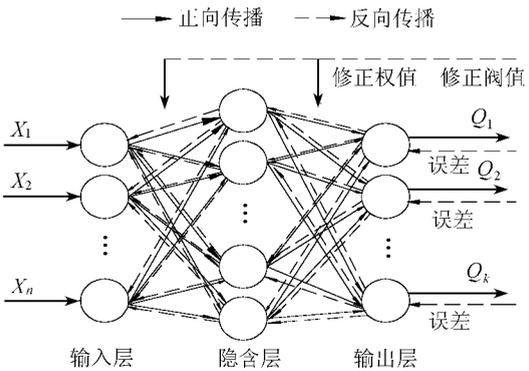


图 2 BP 算法原理示意图

1.2 普通 BP 算法实现步骤

阐述主要计算过程,推导可参阅相关文献^[6-8]。

a. 初始化。设 $x_N = (x_1, x_2, \dots, x_n)$, $y_N = (y_1, y_2, \dots, y_k)$ ($N = 1, 2, \dots, m$) 分别为已归一化的输入、输出样本,其中 m 为样本容量, n 为输入层神经元数。给输入层与隐含层之间的连接权 $\{\omega_{ij}\}$ 和阈值 $\{\theta_j\}$ 及隐含层与输出层之间的连接权 $\{v_{jk}\}$ 和阈值 $\{\gamma_k\}$ 赋予 $(0, 1)$ 之间的随机值。给定学习率 η 、误差精度 ξ 以及最大训练次数。

b. 随机选取一模式对 (x_N, y_N) 提供给网络。

c. 计算隐含层各节点的输入 S_j 、输出 H_j 。

$$H_j = f(S_j) = f\left(\sum_{i=1}^n \omega_{ij}x_i - \theta_j\right) \quad (j = 1, 2, \dots, p) \quad (1)$$

$$f(x) = (1 + e^{-x})^{-1}$$

式中 p 为隐含层神经元数; $f(x)$ 为激励函数。

d. 计算输出层各节点的输入 L_k 、输出 Q_k 。

$$Q_k = f(L_k) = f\left(\sum_{j=1}^p v_{jk}H_j - \gamma_k\right) \quad (k = 1, 2, \dots, q) \quad (2)$$

式中 q 为输出层神经元数。

e. 计算输出层各单元一般化误差 δ_k 。

$$\delta_k = (y_k^k - Q_k)(1 - Q_k)Q_k$$

f. 计算隐含层各单元一般化误差 e_j 。

$$e_j = (1 - H_j)H_j \sum_{k=1}^q \delta_k v_{jk}$$

g. 修正连接权 $\{v_{jk}\}$ 和阈值 $\{\gamma_k\}$

$$v_{jk}(t+1) = v_{jk}(t) + \eta \delta_k^t H_j$$

$$\gamma_k(t+1) = \gamma_k(t) + \eta \delta_k^t$$

h. 修正连接权 $\{\omega_{ij}\}$ 和阈值 $\{\theta_j\}$

$$\omega_{ij}(t+1) = \omega_{ij}(t) + \eta e_j^t x_i$$

$$\theta_j(t+1) = \theta_j(t) + \eta e_j^t$$

i. 随机选择下一模式对提供给网络,返回步骤 c,直到全部 m 个样本训练完毕。

j. 判断全局误差

$$E = \sum_{N=1}^m E_N = \sum_{N=1}^m \sum_k (y_N^k - Q_k)^2 / 2$$

是否小于给定误差 ξ 及训练次数是否达到最大训练次数。如果误差小于给定误差则网络收敛,输出结果,或训练次数达到预定值网络不收敛。否则返回步骤 b,直到全局误差达到给定值。计算流程如图 3 所示。

1.3 算法改进

普通的 BP 神经网络有自身的缺陷:①易陷入局部最小点;②收敛速度慢;③学习过程容易出现震荡。采用学习率自适应和增加动量项相结合的方法对神经网络进行改进^[9]。其权值和阈值调整计算式:

$$u(t+1) = u(t) + \eta(t)D(t) + \alpha[u(t) - u(t-1)] \quad (3)$$

$$\theta(t+1) = \theta(t) + \eta(t)\epsilon(t) \quad (4)$$

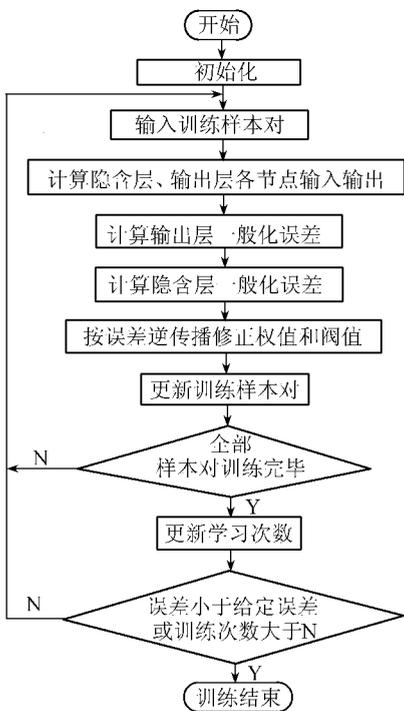


图3 BP网络算法流程

$$D(t) = -\partial E / \partial u(t)$$

式中： $u(t)$ 为权值； $D(t)$ 为 t 时刻的负梯度； $\alpha(t)$ 为阈值； $e(t)$ 为 t 时刻的一般化误差； a 为动量项系数，通常取0.9。加入动量项后，使权重的调整向着底部的平均方向变化，有利于改善网络收敛过程中的震荡，从而改善收敛性。 $\eta(t)$ 为 t 时刻学习率，其随训练过程发生变化，在训练开始阶段，学习率取大些，有利于提高网络的训练速度，在训练的后期，学习率取小些，保证网络收敛在误差最小点。调整算法为

$$\begin{aligned} \eta(t) &= 2^\lambda \eta(t-1) \\ \lambda &= \text{sign}[D(t)D(t-1)] \end{aligned} \quad (5)$$

式(5)表明，当连续两次迭代梯度方向相同时，表明下降太慢，这时可使步长加倍；反之，当连续两次迭代梯度方向相反时，表明下降过头，这时可使步长减半。同时为防止学习率太高而导致震荡或发散，将 η 限制在 $0.01 \sim 0.1^{[5]}$ 。

2 应用实例

2.1 基本资料及其预处理

选取吉林西部26710013号长观井1986~2001年的月平均水位资料进行分析，建立地下水位动态BP网络模型并对模型进行分析。

由于网络输出层的转移函数是单极性Sigmoid函数，它的输出范围是 $[0, 1]$ ；并且Sigmoid函数中间部分对输入的变化比较敏感，而两端对输入的变化反应迟钝，即该函数对取值范围在两端的数据不如

中间部分映射得精确。所以在输入数据之前必须进行样本数据的预处理。

$$T = T_{\min} + \frac{(T_{\max} - T_{\min})(X - X_{\min})}{(X_{\max} - X_{\min})} \quad (6)$$

式中： X 为原始数据； T 为变换后的数据，也称为目标数据； X_{\max} 、 X_{\min} 为原始数据的最大和最小值； T_{\max} 、 T_{\min} 为目标数据的最大和最小值，本文分别取0.8和0.2。

网络运行后，数据按式(6)逆计算还原得到。

2.2 模型的建立

2.2.1 输入、输出样本对的确定

应用BP网络建模进行拟合与预测，必须先确定输入输出样本对。由于本例中地下水位动态变化属于单因素序列，因此先应用相关分析技术，对序列进行自相关识别，在显著性95%水平下，求得自回归阶数为 α （相关系数0.992），即每月水位埋深与前6个月的水位埋深密切相关。所以本文选取前6个月的水位埋深资料作为输入，当月水位埋深作为输出。这样，将1986~2000年资料生成样本对175个，将2001年作为网络检验样本。

2.2.2 网络结构的确定

理论上已证明，三层BP网络能映射任意复杂的非线性函数关系，因此选择三层BP网络。输入层选定6个节点，输出层选为1个节点。隐含层节点数采用“试错法”^[10]确定。初始选定为8个节点，经过多次调试，最终选定为10个节点。亦即网络结构最终确定为6:10:1，即6个输入节点，10个隐含层节点，1个输出节点。

2.3 参数识别

对已建立的模型，应用Visual Basic 6.0根据上述计算步骤和改进算法编制计算程序，该程序在Windows XP系统下运行通过。利用已调试好的程序对网络进行训练，识别参数（即各连接权和阈值）。给定初始学习率为0.1，动量系数0.9，训练精度 ξ 为 3×10^{-4} ，最大训练次数为1万次。运行程序，经1107次训练误差达 2.986×10^{-4} ，小于给定的允许误差，网络收敛，此时学习率为0.01。网络经过学习训练后，学到了预测地下水位动态的知识，这些知识以连接权和阈值的形式储存在网络之中。误差变化见图4。

2.4 精度检验

地下水动态BP网络模型建立后须对其精度进行检验。对上述模型训练结果与实测值进行对比，拟合结果见图5。从图5可看出其拟合精度较高。

经后验差统计检验^[11]，其评价标准^[12]如表1所示。其后验差比值 c 和小误差频率 p 分别为0.09

和 1。所建立的地下水埋深 BP 网络模型的精度达到了“好”的标准。

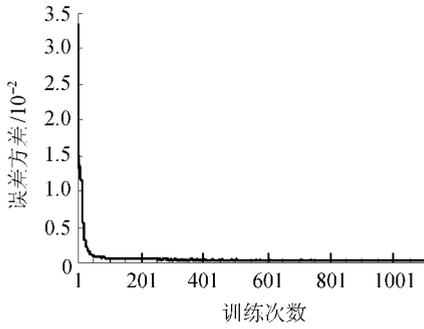


图 4 网络训练误差函数变化曲线

下水位埋深逐月进行预报。水位埋深曲线见图 7。

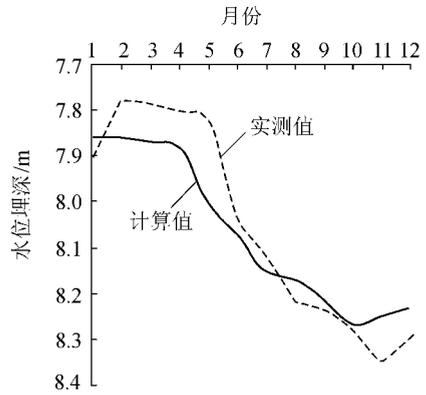


图 6 后验预测拟合曲线

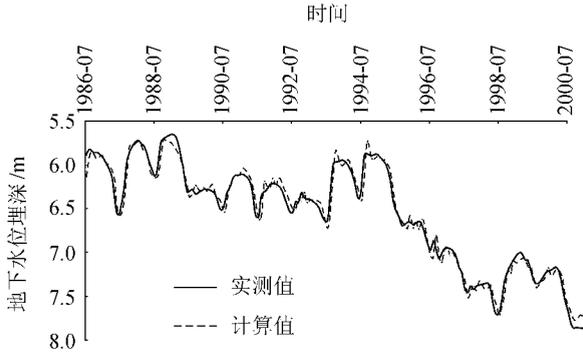


图 5 BP 地下水埋深拟合曲线

表 1 预报精度评价标准

预测等级	p	c	预测等级	p	c
好	> 0.95	< 0.35	合格	> 0.70	< 0.65
良	> 0.80	< 0.50	不合格	≤ 0.70	≥ 0.65

拟合精度高,并不等于模型预报精度也高,因此在将模型用于预报前还要进行预报检验。采用未参加建模的 2001 年水位埋深资料进行后验预测检验。用经过训练得到的权值和阈值,按计算步骤 $a \sim d$ 正向计算得到预测值。其后验预测结果见表 2 和图 6。

表 2 2001 年后验预测误差

月份	绝对误差/m	相对误差/%	月份	绝对误差/m	相对误差/%
1	0.046	0.234	7	-0.031	0.033
2	-0.082	0.683	8	0.049	0.250
3	-0.090	0.785	9	0.018	0.117
4	-0.081	0.668	10	0.007	0.250
5	-0.191	2.032	11	0.104	0.922
6	-0.029	0.016	12	0.069	0.492

从表 2 和图 6 可看出,该模型预测精度较高,绝对误差为 0.007 ~ 0.19 m 之间,最小达到 0.007 m,相对误差均小于 3%,故该模型完全可应用于预报未来地下水埋深。

2.5 模型预报

现依据经过精度检验的模型对 2002 ~ 2006 年地

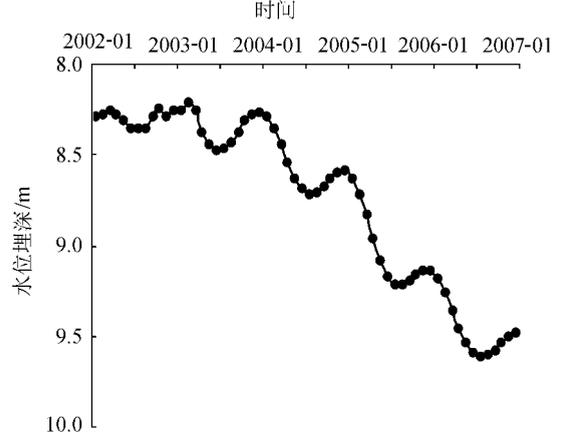


图 7 水位埋深预报

从图 7 中可以看出,如继续按过去的开采模式进行开采,该观测井中地下水位将持续下降,且下降幅度越来越大。若不采取控制地下水超采的有效措施,随着社会经济和工农业的发展,需水量增大,地下水位下降速率和幅度将比预测值更大。经对该区其他观测井的模拟,也可得到相同的结论。因此必须对该区地下水开采利用进行严格控制和科学管理,防止地下水位继续下降而导致该区生态环境恶化。

3 结论

a. 运用自适应学习率动量 BP 算法建立了吉林西部地下水埋深人工神经网络预测模型。实例研究表明,模型拟合精度和预测精度均较高,达到了“好”的标准,说明该模型能够正确描述复杂的地下水系统,揭示地下水系统的非线性特性。且该方法理论简明,计算简单,所需资料较少易获得,可很好的应用于地下水位动态的模拟预测,为区域地下水的合理开发利用及水资源规划提供较为可靠的依据。

b. 在建模过程中,如何正确的确定网络结构是建模的关键,需要反复调试,以便选择最优的模型,使模型能达到最佳拟合和预报精度。

c. 从模型预报结果可看出 (下转第 59 页)

[D].北京 清华大学 ,1996.

- [2] 莫孝翠 杨开 袁德玉. 湖泊内源污染治理中的环保疏浚浅析 [J]. 人民长江 , 2003 , 34(12) : 47-49.
- [3] 田伟君 王超 李勇 等. 城市污染水体强化净化技术研究进展 [J]. 河海大学学报 : 自然科学版 , 2004 , 32(2) : 136-139.
- [4] 颜昌宙 范成新 杨建华 等. 湖泊底泥环保疏浚技术研究展望 [J]. 环境污染与防治 , 2004 , 26(3) : 189-192.
- [5] 杨士健. 骆马湖富营养化发生机制与防治途径初探 [J]. 中国环境监测 , 2004 , 20(3) : 59-62.
- [6] PATEL-SORRENTINO N , MOUNIER S , BENAÏM J Y. Excitation-emission fluorescence matrix to study pH influence on organic matter fluorescence in the Amazon basin rivers [J]. Water Research , 2002 , 36(10) : 2571-2581.
- [7] BAKER A. Fluorescence excitation-emission matrix characterization of river waters impacted by a tissue mill effluent [J]. Environmental Science & Technology , 2002 , 36(7) : 1377-1382.
- [8] WOLFE A P , KAUSAL S S , FULTON J R , et al.

Spectrofluorescence of sediment humic substances and historical changes of lacustrine organic matter provenance in response to atmospheric nutrient enrichment [J]. Environmental Science & Technology , 2002 , 36(15) : 3217-3223.

- [9] LEENHEER J A , CROUE J P. Characterizing aquatic dissolved organic matter [J]. Environmental Science & Technology , 2003 , 37(1) : 19A-26A.
- [10] BAKER A. Fluorescence properties of some farm wastes : Implications for water quality monitoring [J]. Water Research , 2002 , 36(1) : 189-195.
- [11] WU F C , TANOUE E. Isolation and partial characterization of dissolved copper-complexing ligands in streamwaters [J]. Environmental Science & Technology , 2001 , 35(18) : 3646-3652.
- [12] 傅平青. 水环境中的溶解性有机质及其与金属离子的相互作用——荧光光谱学研究 [D]. 北京 : 中国科学院研究生院 , 2004.

(收稿日期 2005-12-19 编辑 舒 建)

(上接第 8 页) 如继续按过去的开采模式进行开采 , 该区地下水位将持续下降 , 且降幅越来越大。应对该区地下水开采利用进行严格控制和科学管理 , 防止地下水位继续下降而导致该区生态环境恶化。

d. 同时应该指出 , 由于 ANN 模型并不是对非线性系统的真实描述 , 不能反映系统内部的真实结构 , 因而不能完全替代系统的机理模型 , 因此如何建立既能反应地下水系统的非线性特征又能描述地下水系统的内部机理的复合模型将是今后 ANN 模型在水科学中的应用研究的主要内容 , 这也将使 ANN 模型的应用前景更加广阔。

参考文献 :

- [1] 周维博. 人工神经网络理论在并渠结合灌区地下水动态预报中的应用 [J]. 西北水资源与水工程 , 2003 , 14(6) : 5-9.
- [2] 刘勇健. 基于智能算法的地下水位动态预测模型的建立和应用 [J]. 水文地质工程地质 , 2004(3) : 55-57.
- [3] 蒋中明 徐卫亚 张新敏. 滑坡地下水位动态预测的径向基函数法 [J]. 岩石力学与工程学报 , 2003 , 22(9) : 1500-1504.

- [4] 胡铁松 袁鹏 丁晶. 人工神经网络在水文水资源中的应用 [J]. 水科学进展 , 1995 , 1(1) : 76-82.
- [5] 屈忠义 陈亚新 史海滨 等. 地下水文预测中 BP 网络的模型结构及其算法探讨 [J]. 水利学报 , 2004(2) : 88-93.
- [6] 金菊良 丁晶. 遗传算法及其在水科学中的应用 [M]. 成都 : 四川大学出版社 , 2000 : 68-88.
- [7] 王伟. 人工神经网络原理-入门与应用 [M]. 北京 : 北京航空航天大学出版社 , 1995 : 52-76.
- [8] 闻新 周露 王丹力 等. Matlab 神经网络应用设计 [M]. 北京 : 科学出版社 , 2000 : 207-212.
- [9] 楼顺天 施阳著. 基于 Matlab 的系统分析与设计-神经网络 [M]. 西安 : 西安电子科技大学出版社 , 1988.
- [10] RAMAN H , SUNILKUM N. Multivariate modeling of water resources time series using artificial neural networks [J]. Hydrological Sciences Journal , 1995 , 40(2) : 145-163.
- [11] 陈南祥 苏万益 吴林娜. 地下水动态的随机模拟模型 [J]. 山西水利科技 , 1994 , 24(1) : 12-17.
- [12] 杨忠平 卢文喜 李平. 时间序列模型在吉林西部地下水动态变化预测中的应用 [J]. 水利学报 , 2005 , 36(12) : 1475-1479.

(收稿日期 2006-02-19 编辑 舒 建)