

DOI :10.3876/j.issn.1000-1980.2010.03.016

基于混沌优化 GMDH 网络的灌区地下水水位预测

朱新国¹ 张展羽¹ 刘 莉²

(1.河海大学水利水电学院,江苏 南京 210098;2.宿迁市水务局,江苏 宿迁 223800)

摘要:为预测灌区地下水水位,提出了一种基于混沌优化的 GMDH 网络预测方法,该方法利用混沌优化算法全局搜索 GMDH 网络的初始权值,并利用优化后的 GMDH 网络建立预测模型对灌区地下水埋深进行预测.算例计算结果表明,该方法能加快 GMDH 网络结构稳定的速度,预测精度较高,可用于对地下水水位的预测.

关键词:地下水;水位预测;GMDH 网络;混沌优化

中图分类号:P641.2 文献标志码:A 文章编号:1000-1980(2010)03-0317-05

灌区地下水埋深动态变化趋势预测对于指导地下水合理开发利用,实现地下水资源的可持续利用,有着重大的理论和现实意义^[1].地下水的动态呈季节性周期性变化,在这种情况下进行水位预测时,一般要将水位原始数据进行处理,分离出周期项、趋势项及随机项,分别进行分析、建模和预测,然后重新组合,才能得到最终结果.传统的地下水分析预测方法主要是确定性数学模型方法和随机统计方法,例如有限差分法、回归分析、时间序列分析、灰色动态模型、神经网络模型等.有限差分法、回归分析等多以线性理论为基础,考虑问题偏于简单化,导致预测精度不高.时间序列分析不能描述预测结果在时序细节上的变化,灰色动态模型、灰色代数等常用的预测方法,很难反映地下水周期性波动的特点.本文首次将混沌优化算法和 GMDH(group method of data handling)神经网络相结合,提出了混沌 GMDH 网络优化算法,并将该算法用于灌区地下水水位预测,探讨该算法的预测精度.

1 混沌优化的 GMDH 网络

1.1 混沌优化算法

混沌(chaos)是自然界广泛存在的一种非线性现象,它貌似随机,却隐含着精致的内在结构,具有遍历性、随机性和对初始条件的极度敏感性,能在一定范围内按其自身规律不重复地遍历所有状态,利用混沌运动的这些性质可以进行优化搜索^[2].Logistic 映射^[3]就是一个典型的混沌系统,迭代公式如下:

$$p_{n+1} = \mu p_n (1 - p_n) \quad (1)$$

式中 μ 是控制参量,当 $\mu = 4$ 时,Logistic 映射系统完全处于混沌状态.利用混沌运动特性可以进行优化搜索,其基本思想是:将混沌状态引入到优化变量中,用类似载波的方法把混沌运动的遍历范围扩大到优化变量的取值范围,利用混沌变量进行搜索.利用混沌的遍历性特点可进行优化搜索且能避免陷入局部极小,这无疑会比随机搜索更具有优越性.更可贵的是,混沌优化不需要知道确切的系统知识,对评价解的优劣情况的目标函数既不要求连续也不要求可微,具有高度的非线性特点.

1.2 GMDH 网络^[4-7]

GMDH 是前馈神经网络中常用于预测的一种网络,其特点是网络结构不是固定的,而是在训练过程中不断改变.图 1 为一个训练后比较经典的 GMDH 网络.

由图 1 可见,GMDH 网络中间层的每个神经元和其前一层的 2 个神经元相对应,输出层的前一层必定只有 2 个神经元.训练从输入层开始构造网络,调整每一个神经元的权值和增加网络层数直到满足映射精度为

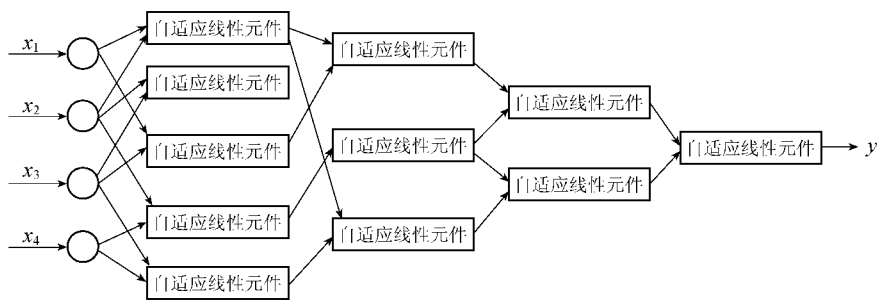


图 1 一个训练过的 GMDH 网络

Fig. 1 A trained GMDH neural network

止. 各层的神经元输入/输出为

$$y_{k,l} = \omega_0 + \omega_1 y_{k-1,i} + \omega_2 y_{k-1,i}^2 + \omega_3 y_{k-1,i} y_{k-1,j} + \omega_4 y_{k-1,j}^2 + \omega_5 y_{k-1,j} \tag{2}$$

式中： $y_{k,l}$ ——第 k 层的第 l 个处理单元，且 $y_{0,l} = x_l$ ； ω_i ($i = 1, 2, 3, 4, 5$)——神经元的权值。因此， $y_{k,l}$ 是权值与输入量的二阶多项式。

采用带非线性二阶预处理器的自适应线性元件 (adaline) 作为神经元，其权值与多项式的系数相对应，同时采用 Widrow-Hoff 学习规则代替线性回归规则，就可将确定多项式系数的工作转换成执行一个改进了的学习程序。假设在时刻 k 神经元的权向量为

$$\boldsymbol{\omega}_k = (\omega_0, \omega_1, \omega_2, \omega_3, \omega_4, \omega_5)^T \tag{3}$$

输入向量为

$$\boldsymbol{x}_k = (1, x_1, x_1^2, x_1 x_2, x_2, x_2^2) \tag{4}$$

由 Widrow-Hoff 学习规则修正：

$$\boldsymbol{\omega}_{k+1}^T = \boldsymbol{\omega}_k^T + \alpha \frac{\boldsymbol{x}_k}{\|\boldsymbol{x}_k\|^2} (y_{dk} - \boldsymbol{\omega}_k^T \boldsymbol{x}_k) \tag{5}$$

式中： y_{dk} ——神经元在时刻 k 的目标输出； α ——学习率，其取值范围为 0.1 ~ 1。按照式 (5) 就可修正 $\boldsymbol{\omega}$ 值，减小网络输出与目标输出之间的差别。

上述计算是在假设网络仅有一个输出的前提下进行的，从式 (5) 可以看出，网络期望输出值 y_{dk} 出现在每个输入层神经元中，并希望通过训练使各神经元都能产生这一期望输出。对于一个神经元，当输入训练数据集中每一个数据后产生的均方差之和 S_E 达到最小时，对这个神经元的训练就结束，其权值予以固定。当输入层的神经元全部训练一遍后，训练停止。这时，选择数据被加到神经元上，并计算相应的 S_E 。那些 S_E 小于阈值的神经元被放入下一层，其余神经元则被舍弃，同时记录每一层神经元训练过程中产生的最小 S_E 。若当前层在训练过程中产生的最小 S_E 小于前一层时，就产生一个新的神经元层，这一层中的神经元数取决于上一层中保留的神经元数，然后对新的神经元层进行训练和选择而保持已训练的神经元层不变，这一过程一直进行到 S_E 不再减小为止，这时取前一层神经元中误差最小的神经元的输出作为网络输出。当新神经元层只有一个神经元且该层的 S_E 小于前一层时，这一神经元就作为输出神经元。输出神经元确定以后，要对网络进行整理，所有与输出神经元无直接或间接联系的神元都被舍去，仅留下那些与输出有关的神元。

1.3 GMDH 网络的混沌优化

神经网络的网络结构、激活函数、训练样本确定以后，网络的误差完全由网络的权值决定^[8]。因此，神经网络的训练过程，也就是寻找一组网络权值使得网络误差最小的过程^[9-11]，故可描述成如 $\min E(\boldsymbol{W})$ 的一个优化问题，其中神经网络权值 $\boldsymbol{W} = (\omega_1, \omega_2, \dots, \omega_l)^T$ 可用平方型的误差函数来表示：

$$E(\boldsymbol{W}) = \frac{1}{2} \sum (d_i - y_i)^2 \tag{6}$$

式中： $E(\boldsymbol{W})$ —— $\boldsymbol{W}$ 对应的网络误差值； d_i ——神经网络的理想输出； y_i ——网络的实际输出。

对于 GMDH 网络，各层的所有神经元都存在对式 (6) 所示的误差函数的优化问题 $\min E(\boldsymbol{W})$ 。对于优化问题 $\min E(\boldsymbol{W})$ ，可采用混沌优化的 GMDH 方法来解决，即先用混沌优化方法得到各神经元初始权值 $\boldsymbol{W}$ ，然后利用 GMDH 网络进行迭代学习。

训练一个用于地下水水位预测的混沌优化的 GMDH 网络的基本步骤如下：

第 1 步 地下水水位数据预处理 ,对数据进行归一化处理.

第 2 步 ,决定网络的输入信号数.对于预测 ,需要用到 n 个过去输出值.如果需要 , n 的值可以通过计算相关系数确定.

第 3 步 ,将地下水水位数据分成训练数据组和选择数据组.

第 4 步 ,建立输入神经元层 ,神经元数与输入地下水水位数据数 m 有关.对每一个输入都有一个神经元与之对应 ,因此相应的神经元数 $d = \frac{1}{2} m(m - 2)$.

第 5 步 混沌优化方法寻优 GMDH 网络初始权值.

a. 随机产生 d 组 $[0, 1]$ 上 6 个随机数 $p_{0,j}, p_{0,j}$ 不能为方程的不动点 ,如果是则加入小扰动 ϵ .

b. 用 Logistic 映射 $p_{i+1,j} = 4p_{i,j}(1 - p_{i,j}) (j = 1, 2, \dots, N, \text{GMDH 网络中 } N = 6)$,得到 d 组 N 个不同的 $p_{i,j}$.

c. 混沌区间 $[0, 1]$ 映射到对应网络初始权值 $x_{i,j}$ 的取值区间：

$$x_{i,j} = p_{i,j}(x_{j\max} - x_{j\min}) + x_{j\min} \tag{7}$$

式中 $x_{i,j}$ 表示 x_j 处于 i 的状态 ,当 $i = 1$ 时 $f^* = f(x^*)$, x^* 为当前最优解 ,计算性能指标：

$$\begin{cases} f_i = f(x_i) \\ x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,N}) \end{cases} \tag{8}$$

d. 计算目标函数值 ,若 $f_i < f^*$,则 $x^* = x_i, f^* = f_i$,如 $\|f^* - f_i\| < \epsilon$,则结束 ,否则 $i = i + 1$,转到 c.

混沌优化中 $x_{i,j}$ 对应每个神经元权值分量 ,寻优得到的一组权值 W^* 作为 GMDH 网络输入层各神经元初始权值.

第 6 步 ,将训练数据组的地下水水位数据作用于输入层的每一个神经元.在时刻 k 取前 m 年地下水水位作为输入信号 ,相应 $m + 1$ 年的地下水水位作为期望输出 y_k ,计算每一神经元的输出误差并修正其权值和均方误差和 ,当均方误差和大于上一循环计算值时 ,训练停止.

第 7 步 ,输入选择数据 ,计算每一神经元的输出均方差.根据差值确定一个阈值 ,选择方差小于阈值 of 的神经元作为下一层神经元.

第 8 步 ,用混沌优化方法得到本层各神经元的初始权值优化值 ,步骤如第 5 步.

第 9 步 ,当本层最小均方差大于前一层神经元的最小均方差或本层仅有一个神经元时 ,停止训练过程.如果训练是由于最小方差偏大而停止的 ,则将前一层神经元作为输出层 ,并重新整理网络 ,转入第 8 步 ;若训练是因本层仅有一个神经元而停止且本次方差小于前一层时 ,则以本层神经元作为输出层并重新整理网络.所谓重新整理就是指舍去那些与输出神经元没有联系的神经元.

第 10 步 ,利用评价数据组检查训练好的网络性能.评价数据组可以是上述训练数据组和选择数据组的结合 ,也可以是一组全新的数据.采用全新数据组可以在更广泛的基础上检查网络性能.

2 混沌优化的 GMDH 网络预测灌区地下水水位算例

选用文献 [12] 中的资料进行实例计算.三江平原创业农场 1985—2006 年地下水水位埋深实测值见表 1.

应用上述混沌优化的 GMDH 网络对地下水水位进行预测.采用 1985—2004 年的数据来建立模型 ,用已建立的模型对 2005 年和 2006 年地下水水位进行外推预测.用 MatLab 语言编程 ,选择四输入端 GMDH 网络 ,即利用前 4 年的数据预测后 1 年的数据 ,对地下水水位进行预测.1985—2004 年的数据共组成 16 组数据样本 ,将前 12 组数据作为训练用 ,后 4 组数据作为选择评价用.训练过程按前述的 10 个步骤依次进行.经过训练 ,地下水水位预测器由两层神经元组成 ,第一层有 2 个神经元 ,第二层有 1 个神经元.拟合和预测的结果见表 1.

由表 1 可知 ,用混沌优化的 GMDH 网络得到的结果与实测值符合较好 ,尤其是 2005 年和 2006 年的数据 ,既未参加训练 ,也未参加选择 ,但预测的结果却很好 ,预测 2005 年地下水位的相对误差仅为 0.765% ,2006 年的为 1.001% .

表 1 三江平原创业农场 1985—2006 年地下水位实测值与计算值

Table 2 Observed and calculated values of groundwater levels of Chuangye Farm on Sanjiang Plain from 1985 to 2006							
年份	实测值/m	混沌优化 GMDH 网络计算值/m	相对误差/%	年份	实测值/m	混沌优化 GMDH 网络计算值/m	相对误差/%
1985	3.41			1996	5.80	5.86	1.035
1986	3.58			1997	6.08	6.06	0.329
1987	3.75			1998	7.03	7.02	0.142
1988	3.91			1999	8.57	8.56	0.117
1989	3.96	3.98	0.505	2000	9.54	9.55	0.105
1990	4.24	4.25	0.236	2001	10.11	10.13	0.198
1991	4.58	4.59	0.218	2002	10.65	10.64	0.094
1992	4.85	4.88	0.619	2003	11.18	11.16	0.179
1993	5.22	5.24	0.383	2004	11.50	11.49	0.087
1994	5.46	5.42	0.733	2005	11.76	11.85	0.765
1995	5.68	5.66	0.352	2006	11.99	12.11	1.001

GMDH 在建立网络模型时,通过在样本拟合精度和新数据集预测精度之间寻找平衡点,确保了即使是在小样本或数据噪声较大时,算法仍能最大限度上反映系统真实的内部关系,从而确保了所建模型的最优性和泛化能力。

3 结 语

基于 GMDH 网络和混沌优化互补性提出的混沌优化的 GMDH 网络算法无论是学习速度与精度,还是预测结果的可靠性方面都有较大改进,有效地提高了 GMDH 网络技术应用于地下水水位预测的有效性和泛化能力。神经网络是一种有效的非线性智能理论,它在解决地下水水位预测建模方面具有较强的优势,但应该指出,要想准确科学地预测地下水位,除了预测方法之外,如何进一步加强地下水位动态影响因素的研究及地下水动态长期观测,以使该理论的应用能有更完备可靠的基础数据也是一个亟待加强的工作。

参考文献:

[1] 朱长军,郝振纯,周继红,等.组合灰色神经网络法在地下水动态预测中应用[J].辽宁工程技术大学学报:自然科学版,2005,24(增刊2):260-262.(ZHU Chang-jun,HAO Zhen-chun,ZHOU Ji-hong,et al. Application of combined gray neural network method in dynamic prediction of groundwater level[J]. Journal of Liaoning Technical University :Natural Science Edition,2005,24(Sup2) 260-262.(in Chinese))

[2] 张建雄,唐万生.基于混沌遗传算法的一类非线性两层混合整数规划问题求解[J].系统工程理论方法应用,2005,14(5):429-433.(ZHANG Jian-xiong,TANG Wan-sheng. Chaos genetic algorithm method for a class of nonlinear bilevel mixed integer-programming problem[J]. Systems Engineering-Theory Methodology Applications,2005,14(5) 429-433.(in Chinese))

[3] 李兵,蒋慰孙.混沌优化方法及其应用[J].控制理论与应用,1997,14(4):613-615.(LI Bing,JIANG Wei-sun. Chaos optimization algorithm and its application[J]. Control Theory and Applications,1997,14(4) 613-615.(in Chinese))

[4] 肖婷,黄国建.GMDH 神经网络在数据预测中的应用[J].华东船舶工业学院学报,2000,14(4):72-76.(XIAO Ting,HUANG Guo-jian. GMDH neural network application to prediction[J]. Journal of East China Shipbuilding Institute,2000,14(4):72-76.(in Chinese))

[5] KOZUBOVSKIY S F. Use of GMDH to optimize the basic parameters of a jet system for measuring velocities of air streams[J]. Soviet Journal of Automation and Information Science,1987,20(1) 85-90.

[6] 胡明星,郭玲香,郭达志.地下水水位动态预测新方法:数据处理组合法[J].煤田地质与勘探,1998,26(5):38-41.(HU Ming-xing,GUO Ling-xiang,GUO Da-zhi. A new method to predict groundwater level :GMDH[J]. Coal Geology & Exploration,1998,26(5) 38-41.(in Chinese))

[7] 陈建平,杨宜民,张会章,等.一种基于 GMDH 模型的神经网络学习算法[J].云南大学学报:自然科学版,2008,30(6):569-574.(CHEN Jian-ping,YANG Yi-min,ZHANG Hui-zhang,et al. A neural network learning algorithm based on GMDH mode[J]. Journal of Yunnan University :Natural Sciences,2008,30(6) 569-574.(in Chinese))

[8] 唐巍,张学义,李殿璞.神经网络权值的混沌优化方法研究[J].哈尔滨工程大学学报,2000,21(3):12-14.(TANG Wei,ZHANG Xue-yi,LI Dian-pu. Chaotic optimization method for neural network's weights[J]. Journal of Harbin Engineering University,2000,21(3) 12-14.(in Chinese))

[9] 刘勇健 ,张伯友 . 基于混沌优化-神经网络的水泥搅拌桩复合地基承载力预测方法[J]. 路基工程 ,2006(1) :63-66.(LIU Yong-jian ,ZHANG Bo-you . On account of choas optimization-neural network for predication bearing capacity on composite foundation of deep mixing pile[J]. Subgrade Engineering ,2006(1) :63-66.(in Chinese))

[10] 李红霞 ,张建雄 ,赵新华 . 基于混沌优化 BP 神经网络的地面沉降模型[J]. 中国矿业大学学报 ,2008 ,37(3) :396-401.(LI Hong-xia ,ZHANG Jian-xiong ,ZHAO Xin-hua . A land subsidence model based on a chaos optimization algorithm and a BP neural network [J]. Journal of China University of Mining & Technology ,2008 ,37(3) :396-401.(in Chinese))

[11] 尹光志 ,谭钦文 ,魏作安 . 基于混沌优化神经网络的冲击地压预测模型[J]. 煤炭学报 ,2008 ,33(8) :871-875.(YIN Guang-zhi ,TAN Qin-wen ,WEI Zuo-an . Combined optimization model of rock-burst prediction based on chaos optimization and neural networks[J]. Journal of China Coal Society ,2008 ,33(8) :871-875.(in Chinese))

[12] 盖兆梅 ,付强 ,刘仁涛 . 基于混沌粒子群优化算法的灰色 GM(1 ,1)模型在地下水埋深预测中的应用[J]. 数学的实践与认识 ,2008 ,38(11) :67-72.(GAI Zhao-mei ,FU Qiang ,LIU Ren-tao . Grey GM(1 ,1) model based on chaotic particle swarm optimization algorithm and its application in groundwater dynamic prediction[J]. Mathematics in Practice and Theory ,2008 ,38(11) :67-72.(in Chinese))

Prediction of groundwater levels in irrigation districts based on
chaos optimization of GMDH neural network

ZHU Xin-guo¹ , ZHANG Zhan-yu¹ , LIU Li²

(1. College of Water Conservancy and Hydropower Engineering , Hohai University , Nanjing 210098 , China ;
2. Suqian Water Affairs Bureau , Suqian 223800 , China)

Abstract : A prediction method based on the chaos optimization of group method of data handling (GMDH) neural network is put forward for the groundwater levels of irrigation districts . The chaos optimization algorithm was employed to globally search the initial weight values of the GMDH network . By using the optimized GMDH network , the relevant model was established to predict the groundwater levels of irrigation districts . The calculated results of cases show that the proposed method can accelerate the stabilizing of the GMDH network structure , that it has high predictive accuracy , and that it is of practical value for the prediction of groundwater levels .

Key words : groundwater ; level prediction ; GMDH neural network ; chaos optimization