

文章编号:1001-4179(2011)19-0077-03

基于粒子群神经网络的凌汛开河日期预测研究

赵晓慎, 吴海波, 王文川, 李倩

(华北水利水电学院 水利学院, 河南 郑州 450011)

摘要:为了更准确地预测凌汛开河日期,提出用粒子群算法和 BP 神经网络相结合的粒子群神经网络模型。介绍了模型的设计和算法实现的流程。该模型通过粒子群算法对 BP 神经网络初始的权值和阈值进行优化,并以黄河内蒙段三湖河口站作为研究实例进行冰凌开河日期预测。结果表明,经粒子群优化后的 BP 神经网络预测精度比遗传神经网络和单一 BP 神经网络更高。

关键词:凌汛; 开河日期; 粒子群优化算法; BP 神经网络; 遗传算法

中图分类号: P64 **文献标志码:** A

凌汛是高纬度地区冬季特有的一种现象,不但严重影响冬季的水力发电、航运交通、水利施工,而且还对沿岸群众的生命和财产安全带来极大的危害^[1-2]。影响河流冰情的因素主要包括热力因素、动力因素、河势因素,但各个因素与冰情之间的确切函数关系有待进一步研究^[2-3]。目前,国内外对封河、开河日期预报模型较多,H. T. Shen 等依据热交换原理和冰水力学理论建立了数学模型^[4-5],陈守煜等利用传统的 BP 神经网络及遗传改进算法对凌汛开河和封河日期进行了预报^[1-3,6-8],李亚伟等依据统计学理论建立了凌汛支持向量机预测模型及冰情投影寻踪回归模型^[9-10],均取得了很好的预报效果。BP 神经网络具有大规模的分布式存储与并行运算能力、非线性动力学映射能力、较强的鲁棒性等优点,同时也存在网络结构选择困难、容易陷入局部最优值点和阈值初值的选取等问题^[11-12]。

为了克服 BP 神经网络在预测中的缺陷,本文采用粒子群算法对 BP 神经网络初始的权值和阈值进行优化,使改进后的 BP 神经网络能够更好地预测输出,并将其应用于黄河内蒙段的三湖河口站凌汛开河日期的预测,实例表明,改进后的 BP 网络预测效果更佳。

1 粒子群神经网络模型简介

1.1 粒子群优化算法

粒子群优化算法(PSO)由 J. Eberhart 和 R. C. Kennedy 于 1995 年提出^[13],是通过模拟鸟群随机觅食过程中个体之间相互协作使集体达到最优的一种全局寻优算法。该算法中每个粒子都代表问题的一个潜在解,用位置、速度和适应值 3 项指标表示该粒子特征。粒子的速度决定了粒子移动的方向和距离,适应度值由适应度函数计算得到,其值的好坏表示粒子的优劣。粒子在解空间寻优迭代过程中,通过跟踪个体极值 Pbest 和群体极值 Gbest 更新个体自身的位置,其中个体极值 Pbest 是指个体所经历位置中计算得到的适应度值最优位置;群体极值 Pbest 是指种群中的所有粒子搜索到的适应度最优位置。粒子每迭代一次,就计算一次适应度值,通过自身不断与更新的粒子在适应度值和个体极值、群体极值的适应度值进行比较后,再更新个体极值 Pbest 和群体极值 Gbest 位置,直到迭代结束。

设有 n 个粒子组成的种群 $X = (X_1, X_2, \dots, X_n)$, 粒子 i 的信息用 D 维向量表示,其中第 i 个粒子表示为

收稿日期:2011-04-16

基金项目:河南省教育厅自然科学研究计划项目(2010B570002);华北水利水电学院高层次人才科研启动资助项目(200821)

作者简介:赵晓慎,女,副教授,硕士生导师,主要从事水文水资源等方面的教学与研究工作。E-mail:zhaoxiaoshen@ncwu.edu.cn

$X_i = (x_{i1}, x_{i2}, \dots, x_{id}), i = 1, 2, \dots, n$ 。根据选入的目标函数可计算出每个粒子 X_i 对应的适应度值^[14-15]。记第 i 个粒子的速度为 $V_i = (V_{i1}, V_{i2}, \dots, V_{id})$, 其个体极值为 $P_i = (p_{i1}, p_{i2}, \dots, p_{id})$, 整个粒子群的全局极值为 $P_g = (p_{g1}, p_{g2}, \dots, p_{gd})$ 。在每次迭代过程中, 粒子通过个体最佳位置和群体极值, 并根据下式更新自身的速度和位置:

$$\left. \begin{aligned} V_{id}^{l+1} &= \lambda V_{id}^l + \kappa_1 \varphi_1 (P_{id}^l - X_{id}^l) + \kappa_2 \varphi_2 (P_{gd}^l - X_{id}^l) \\ X_{id}^{l+1} &= X_{id}^l + V_{id}^{l+1} \end{aligned} \right\} (1)$$

$$(d = 1, 2, \dots, D; i = 1, 2, \dots, n)$$

式中, λ 为非负常数, 称为惯性因子; l 为当前迭代次数; V_{id} 为粒子的速度, $V_{id} \in [-V_{\max}, V_{\max}]$, V_{\max} 为常数; κ_1 和 κ_2 为非负常数, 称为加速度因子; φ_1 和 φ_2 为分布于 $[0, 1]$ 之间的随机数。为了防止粒子的盲目搜索, 一般迭代终止条件为最大迭代次数或是粒子群迄今为止搜索到的最优极值满足适应阈值。

1.2 粒子群优化 BP 神经网络模型设计

采用粒子群算法与 BP 神经网络相结合, 通过定义粒子群算法和 BP 神经网络相关参数, 构造一个 3 层的粒子群神经网络组合模型。当网络结构确定后, 需要对网络连接的权值和阈值用粒子群算法进行优化, 通过粒子个体位置的不断更新, 使 BP 神经网络的预测输出和期望输出的均方误差 (适应度值) 减小到允许的精度, 采用下式计算粒子个体的适应性度值:

$$E = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i^d - y_i)^2} \quad (2)$$

式中, N 为训练的样本数; y_i^d 为第 i 个样本的理想输出值; y_i 为第 i 个样本的期望输出值。

2 粒子群算法优化 BP 神经网络的实现

粒子群算法优化 BP 神经网络与遗传算法优化 BP 神经网络构架一样, 分为 BP 神经网络结构确定、粒子群算法优化过程和 BP 神经网络预测 3 个部分^[16]。BP 神经网络结构根据拟合函数输入输出参数的个数设定, 进而确定粒子群算法中个体的长度。文中粒子群算法优化过程是指对 BP 神经网络的权值和阈值进行优化选取, 即粒子种群中的每个个体都包含了一个网络所需的所有权值和阈值, 每个个体都通过适应性函数计算其适应度值, 并对种群中的个体进行若干次的速度和位置更新得到最优个体适应值。BP 神经网络预测是将最优个体所包含的权值和阈值赋给 BP 神经网络, 使网络经训练、仿真后得到预测输出。粒子群优化 BP 神经网络相关步骤如下。

(1) 根据 BP 神经网络输入输出变量确定其结构, 且定义粒子群算法控制参数, 对初始粒子位置和粒子速度赋予随机值。

(2) 对输入输出变量的数据进行预处理和 BP 神经网络训练, 通过预测输出和期望输出之间的均值误差形成适应函数, 且计算每个初始粒子适应度值确定个体极值和群体极值。

(3) 根据式 (1) 更新粒子速度和位置, 计算新粒子的适应度值, 更新个体的极值和群体极值, 将更新得到的最优个体对 BP 神经网络赋初始的权值和阈值, 直到满足条件为止, 否则继续进行个体更新。

(4) 提取最优个体将值赋给 BP 神经网络作为权值和阈值, 再进行预测输出。其整个步骤流程如图 1 所示。

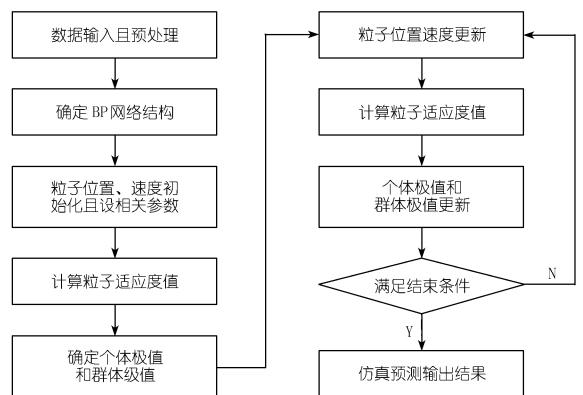


图 1 粒子群神经网络算法流程

3 实例应用

3.1 数据源和参数设置

影响凌汛的因素较多, 可概括为热力因素、动力因素及河势因素, 其中又以热力因素和动力因素影响最大, 包括气温、水温、降水、流量、流速、风速、水位等因子^[1-2, 17]。根据凌汛产生机理, 文中以黄河内蒙段三湖河口站采集的水位、累计正气温、流量、封冻期最大冰厚及开河历时 5 个因子的凌汛数据进行分析^[2], 其中以 1968 ~ 1996 年的数据为样本数据, 后续 5 a 的数据为测试数据。利用前 4 个因子作为 BP 神经网络的输入变量, 以开河历时因子为网络输出变量。据 Kolmogorov 定理设定 BP 神经网络结构的隐含层有 6 个神经元^[18], 因此构造 3 层 BP 神经网络结构为 $4 \times 6 \times 1$, 取迭代次数 $epochs = 200$, 学习率 $lr = 0.15$, 目标函数误差值 $goal = 0$ 。粒子群算法在优化过程中取种群规模 $sizepop = 200$, 最大迭代次数 $maxgen = 50$, 惯性因子 $\lambda = 1$, 加速因子 $\kappa_1 = 2$, $\kappa_2 = 1.8$ 。为了防止粒子

的盲目搜索,提高粒子群运行速度,一般将粒子的位置和速度限制在一定的范围内,设粒子个体最大位置 $X_{\max} = 3$ 及最大速度 $V_{\max} = 5$ 。

3.2 预报结果比较

以开河历长为粒子群神经网络输出变量,其余为输入变量。本文选用权值和阈值未经优化的 BP 神经网络模型与遗传算法优化 BP 神经网络作比较。遗传算法的种群规模为 200,最大迭代次数为 50,种群个体范围 $[-3, 3]$,其余参数设为默认值。经过运算,得到 3 种预测模型的误差及与实际值的比较见表 1。可以看出粒子群神经网络较遗传神经网络和 BP 神经网络有更强的预测能力。通过对表 1 预测结果的分析,粒子群神经网络预测结果的平均相对误差精度比遗传神经网络模型和单一 BP 神经网络模型提高了 4.45% 和 6.94%,误差减少了 4.82% 和 7.72%,与文献[2]中模糊神经网络预测值比较,误差减少了 5.53%。

表 1 3 种模型预测结果比较 d

年份	实测值	GA - BP		BP		PSO - BP	
		预测值	相差天数	预测值	相差天数	预测值	相差天数
1996 ~ 1997	42	37	5	42	0	42	0
1997 ~ 1998	40	38	2	41	1	41	1
1998 ~ 1999	44	46	2	36	8	42	2
1999 ~ 2000	40	34	6	49	9	42	2
2000 ~ 2001	43	42	1	41	2	42	1

注:GA - BP 为遗传神经网络;PSO - BP 为粒子群神经网络。

4 结语

通过对凌汛开河日期变化规律进行分析,利用粒子群算法对 BP 神经网络初始的权值和阈值进行全局优化,提出凌汛预测的粒子群神经网络模型。该模型成功克服了 BP 神经网络在预测中极易陷入局部极小点等缺点,并将模型具体运用于黄河冰凌开河日期的预测,其结果与实际值更加接近。粒子群神经网络是

一种预测效果较好的群智能优化预测技术,对防凌及其他受多因素影响的问题预测工作具有现实意义。

参考文献:

- [1] 王志兴,李成振,范宝山,等.改进的遗传 BP 交叉训练算法及其在河流冰清预报中的应用[J].水力发电学报,2010,29(1):76-79.
- [2] 陈守煜,冀鸿兰.冰凌预报模糊优选神经网络 BP 方法[J].水利学报,2004,(6):114-118.
- [3] 宋安,王世杰,卢海.基于 BP 神经网络的流凌开河日期预报模型应用[J].天津大学学报,2008,41(6):731-735.
- [4] Shen H T. Dynamic transport of river ice[J]. Journal of Hydraulic Research,1990,28(6):659-671.
- [5] Shen H T. Under cover transport and accumulation of frazil granules [J]. Journal of Hydraulic Engineering,1995,121(2):184-195.
- [6] 王涛,杨开林,郭永鑫,等.神经网络理论在黄河宁夏河段冰清预报中的应用[J].水利学报,2005,36(10):1204-1208.
- [7] 王志兴,李成振,范宝山,等.基于遗传神经网络的河流冰凌预报[J].水利水电技术,2009,40(2):57-59.
- [8] 卜松,王军,郭立文. ACA - BP 网络再冰塞水位预测中的应用[J].人民黄河,2009,31(5):45-47.
- [9] 李亚伟,陈守煜,韩小军.基于支持向量机 SVR 的黄河凌汛预报方法[J].大连理工大学学报,2006,46(2):272-275.
- [10] 王兴志,李成振,陈刚.冰情预报的投影跟踪回归模型[J].自然灾害学报,2009,18(5):174-177.
- [11] 岳琳,张宏伟,王亮.粒子群优化算法在城市需水量预测中的应用[J].天津大学学报,2007,40(6):742-746.
- [12] 刘洪波,王秀坤,孟军.神经网络基于粒子群优化的学习算法研究[J].小型微型计算机系统,2005,26(4):638-640.
- [13] Kennedy J, Eberhart R C. Particle swarm optimization[C]// Proc. IEEE Int'l Conf. on Neural Networks. Perth: [s. n.], 1995.
- [14] 杨道辉,马光文,刘起方,等.基于粒子群优化算法的 BP 网络模型在径流预测中的应用[J].水力发电学报,2006,25(2):65-68.
- [15] 周品,赵新芬. MATLAB 数学建模与仿真[M].北京:国防工业出版社,2009.
- [16] MATLAB 中文论坛. MATLAB 神经网络 30 个案例[M].北京:北京航空航天大学出版社,2010.
- [17] 陆钦年,段忠东,欧进萍,等.黑龙江省的凌汛灾害及其区域划分[J].自然灾害学报,2000,9(3):54-58.
- [18] 葛哲学,孙志强.神经网络理论与 MATLABR2007 实现[M].北京:电子工业出版社,2007.

(编辑:李慧)

Prediction of beginning date of ice jam flood based on particle swarm optimization - neural network model

ZHAO Xiaoshen, WU Haibo, WANG Wenchuan, LI Qian

(College of Water Conservancy, North China University of Water Resources and Electric Power, Zhengzhou 450011, China)

Abstract: In order to predict the beginning date of ice jam flood more accurately, the particle swarm optimization - neural network model is proposed by the combination of particle swarm optimization algorithm and BP neural network. The design of the model and flow of the algorithm realization is presented. The initial weights and thresholds of BP network are optimized by using particle swarm optimization algorithm, and taking the Sanhuhekou Station of Yellow River in the Inner Mongolia as example, the beginning date of ice jam flood is predicted by the proposed model. The results show that the BP neural network optimized by particle swarm algorithm has higher prediction precision than the genetic algorithm and BP neural network.

Key words: ice jam flood; beginning date; particle swarm optimization algorithm; BP neural network; genetic algorithm