

径流中长期预报模糊优选神经网络模型应用研究

刘艳丽, 袁晶瑄, 周惠成*

(大连理工大学 土木水利学院, 辽宁 大连 116024)

摘要: 预报因子选择与模型训练精度确定, 是模糊优选神经网络模型应用于径流中长期预报时有待研究解决的两个重要问题. 应用预报因子集与预报量间的复合非线性相关分析方法选择预报因子(集), 克服了通常单因子线性相关分析选择预报因子的不适用性; 通过定义综合效应系数来综合评价模糊优选神经网络模型的拟合能力与外推预报能力, 为研究模型的拟合精度高而外推预测精度低的问题提供了一种解决方法.

关键词: 模糊优选; 神经网络; 年径流预报; 预报因子选择; 综合效应系数

中图分类号: P33; TV124 **文献标志码:** A

0 引言

径流中长期预报作为中长期水文预报的重要组成部分, 对水资源有效利用、水力发电、水库调度等的指导意义, 越来越受到学者们的关注. 作为流域的水文气象、自然地理、地质地貌、植被覆盖等诸多因子综合作用的结果, 年径流过程是一弱相依、高维非线性复杂动力系统, 国内外的年径流预测研究至今仍处于探索阶段^[1,2]. 现有的年径流预测方法可大致分为成因预测方法和统计预测方法两大类. 由于年径流具有时间和空间上的复杂统计特性, 目前成因预测因受中长期天气预报水平的制约仍极为困难; 统计预测方法包括历史演变法、时间序列分析法、多元(逐步)线性回归法, 这些方法的预报成果目前仍难以有效地指导生产实践, 需要进一步研究新的方法. 人工神经网络方法因其非线性模拟的优势在复杂的水文系统预报中发挥了很大作用^[3~5], 但存在拟合精度高而预测精度低的现象. 从预报机理上研究中长期预报, 结合描述非线性系统较好的神经网络模型, 将是水文中长期预报的一个积极探索.

预报因子的选择是预报的关键步骤之一, 目前多用单因子与预报量间的线性相关系数加以显著性检验^[6,7], 但径流中长期预报由于其复合非线性特点, 单因子的线性相关系数将不再适用.

本文将陈守煜提出的模糊优选神经网络模型^[8]应用到年径流预测中, 针对预报因子与预报量之间的复杂非线性关系, 利用复合非线性系数选择预报因子; 对过拟合问题, 定义综合效应系数来综合反映模型拟合精度和预测精度, 对模型进行优化.

1 模糊优选神经网络预报模型^[8]

设预报量的样本集 $\mathbf{Y} = (y_1 \ y_2 \ \cdots \ y_n)$, 筛选出 m 个预报因子, 则其特征值矩阵为

$$\mathbf{X} = (x_{ij})_{m \times n} \quad (1)$$

式中: x_{ij} 为样本 j 预报因子 i 的特征值, $i = 1, 2, \cdots, m; j = 1, 2, \cdots, n$. 当样本预报量与预报因子 i 正相关时, 相对隶属度由式(2)确定:

$$r_{ij} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

式中: $x_{\min}$ 为预报因子 i 的样本最小值; $x_{\max}$ 为预报因子 i 的样本最大值.

当样本预报量与预报因子 i 负相关时, 相对隶属度由式(3)确定:

$$r_{ij} = \frac{x_{\max} - x_{ij}}{x_{\max} - x_{\min}} \quad (3)$$

根据矩阵 $\mathbf{X}$ 用式(2)、(3)得到预报因子对“丰”的相对优属度矩阵 $\mathbf{R} = (r_{ij})_{m \times n}$.

预报量 $\mathbf{Y}$ 对“丰”的相对隶属度公式为

收稿日期: 2006-03-01; 修回日期: 2008-03-20.

基金项目: 国家自然科学基金委员会、二滩水电开发有限公司雅砻江水电联合研究基金资助项目(50579095).

作者简介: 刘艳丽(1981-), 女, 博士生, E-mail: liuyli@student.dlut.edu.cn; 周惠成*(1958-), 男, 教授, 博士生导师.

$$a_j = \frac{y_j - y_{\min}}{y_{\max} - y_{\min}} \quad (4)$$

式中: $y_{\min}$ 为预报量样本最小值; $y_{\max}$ 为预报量样本最大值.

从而得到预报量规格化矩阵

$$\mathbf{A} = (a_1 \quad a_2 \quad \cdots \quad a_n)$$

神经网络模型如图1所示. 其是一个3层的神经网络系统, 输入层有 m 个输入节点, 即有 m 个预报因子; 隐含层有 L 个隐节点, 即有 L 个单元系统, 输出层仅有1个单节点输出.

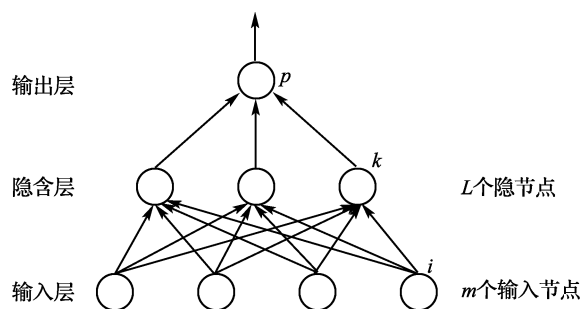


图1 三层神经网络模型

Fig. 1 Three layers neural network architecture

在输入层节点 i 将信息直接传给隐含层节点, 故节点输出与输入相等, 即

$$u_{ij} = r_{ij} \quad (5)$$

对隐含层节点 k , 其输入为

$$I_{kj} = \sum_{i=1}^m \omega_{ik} r_{ij} \quad (6)$$

输出为

$$u_{kj} = 1 / \left\{ 1 + \left[\left(\sum_{i=1}^m \omega_{ik} r_{ij} \right)^{-1} - 1 \right]^2 \right\} = 1 / [1 + (I_{kj}^{-1} - 1)^2] \quad (7)$$

式中: ω_{ik} 为输入层 i 与隐含层 k 的连接权重.

输出层仅一个节点 p , 输入为

$$I_{pj} = \sum_{k=1}^L \omega_{kp} u_{kj} \quad (8)$$

式中: ω_{kp} 为隐含层 k 与输出层节点 p 的连接权重.

输出为

$$u_{pj} = 1 / \left\{ 1 + \left[\left(\sum_{k=1}^L \omega_{kp} u_{kj} \right)^{-1} - 1 \right]^2 \right\} = 1 / [1 + (I_{pj}^{-1} - 1)^2] \quad (9)$$

网络的实际输出 u_{pj} 就是模糊优选神经网络对输入 r_{ij} 的响应, 设样本的期望输出为 $M(u_{pj})$, 其迭代误差为

$$\epsilon = \frac{1}{n} \sum_{j=1}^n [u_{pj} - M(u_{pj})]^2 / 2 \quad (10)$$

2 预报因子的选择及模型参数识别

通常, 以单个预报因子与预报量间的线性相关系数评价预报因子的有效性, 决定是否在预报模型中引入某一因子. 很显然, 由于预报因子和预报量间的非线性复合关系, 仅以单个因子与预报量间的相关关系来决定因子取舍是不合适的. 模型的输出结果是参与预报因子集与预报量之间的复杂非线性相关关系(简称复合非线性相关关系)的最终体现. 利用模型输出与预报量间的复合非线性相关关系来评价预报因子集的有效性, 会与模型本身相协调.

2.1 预报因子集与预报量间的复合非线性分析方法

设定模型训练样本的迭代误差, 对原预报因子集和加入新预报因子的新因子集, 计算比较两次预报结果的复合非线性相关系数差值 (D_0 代表原模型复合非线性相关系数, D_1 代表加入新因子后的模型复合非线性相关系数), 若 $D_1 > D_0$, 则认为新因子加入后对模型有所改善, 建议引入, 但可视改进程度与模型结构的复杂程度而综合确定; 若 $D_1 \leq D_0$, 则认为新加入的因子对模型没有改进, 不予引入.

洪水预报的确定性系数^[9] 反映了预报过程与实际过程的拟合程度, 本文将定义为预报值与实测值之间的复合非线性相关系数, 即

$$D = 1 - \frac{\sum_{i=1}^n [y_c(i) - y_0(i)]^2}{\sum_{i=1}^n [y_0(i) - \bar{y}_0]^2} \quad (11)$$

式中: D 为复合非线性相关系数; $y_0(i)$ 为实测值; $y_c(i)$ 为预报值; $\bar{y}_0$ 为实测值的均值; n 为资料序列长度.

2.2 预报模型训练参数——允许迭代误差确定

模糊优选神经网络模型训练中的一个重要参数是训练样本的允许迭代误差. 允许迭代误差表面上是在控制模型的拟合精度, 但实质上对同一模型结构, 其参数是由迭代误差所决定的. 一方面, 只要因子足够多、模型结构足够复杂, 理论上可使神经网络模型达到任意拟合精度^[10]; 但另一方面, 存在着模型拟合能力高而预测能力低的突出问题, 即过度拟合问题. 本文定义综合效应系数 μ 来综合反映模型的拟合和预报能力, 即

$$\mu = \alpha_1 D_1 + \alpha_2 D_2 = f(\epsilon) \quad (12)$$

式中:定义 $\alpha_1 = \frac{e_1}{e_1 + e_2}$ 、 $\alpha_2 = \frac{e_2}{e_1 + e_2}$ 分别为训练与检验样本精度的权重,满足 $\alpha_1 + \alpha_2 = 1$; D_1 、 D_2 分别为训练与检验样本的复合非线性相关系数; e_1 、 e_2 分别为训练与检验样本的预报结果相对误差均值。

μ 通过 α_1 与 α_2 寻求模型拟合与外推预报能力的均衡。由于 μ 是迭代误差 ϵ 的函数,可通过优化综合效应系数 μ 确定模型训练的迭代误差 ϵ ,进而得到相应的神经网络模型的网络连接权重等参数。

3 径流长期预报实例研究

辽宁桓仁水库是浑江梯级开发的龙头,流域面积 10 364 km²,占浑江流域面积的 70%,工程以发电为主。本文以桓仁水库流域的年径流预报为例,研究模糊优选神经网络模型在径流中长期预报中的应用,其中重点分析预报因子的选择与预报模型参数识别的方法。

3.1 预报因子的评价与选择

3.1.1 预报因子的初步选定 结合桓仁水库已有的大量预报实践及其资料情况,初步选取 7 个前期预报因子,即预报因子 x_1 为前一年 11 月至当年 2 月桓仁水库的总降雨量,预报因子 x_2 为当年 3 月桓仁水库的总降雨量,预报因子 x_3 为前一年 8 月欧亚地区月平均纬向环流指数,预报因子 x_4 为前一年 5 月欧亚地区月平均径向环流指数,预报因子 x_5 为前一年 3 月至 5 月桓仁水库的总降雨量,预报因子 x_6 为前一年 12 月欧亚地区月平均纬向环流指数,预报因子 x_7 为桓仁水库前一年年径流。其中年径流指水文年,即 04-01 ~ 03-31 的年径流。

资料长度为 49 a(1954~2002 年),各因子的变化如图 2、3 所示。将前 31 a 资料作为网络的训练样本,中间 10 a 作为检验样本,与前 31 a 配合、综合考虑模型的拟合与检验精度确定模型参数;后 8 a 不参加确定模型参数,纯粹用于检验所确定模型及其参数的外推预报能力。所分的 3 段资料均包括了丰、平、枯年份,具有一定的代表性。

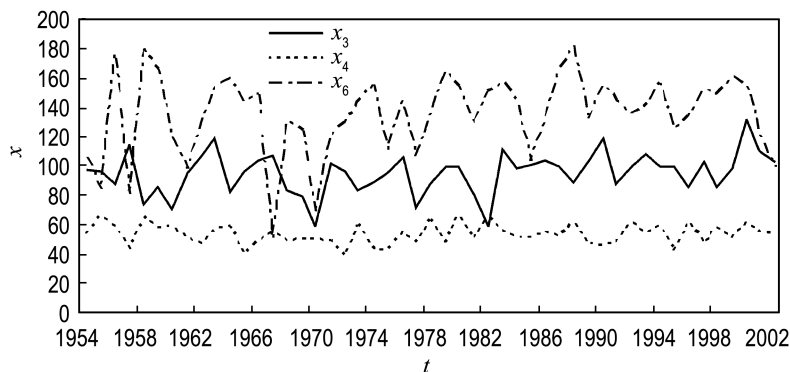


图 2 桓仁水库年径流量的 3 个预报因子特征值

Fig. 2 Other three forecast factors of annual runoff of Huanren Reservoir

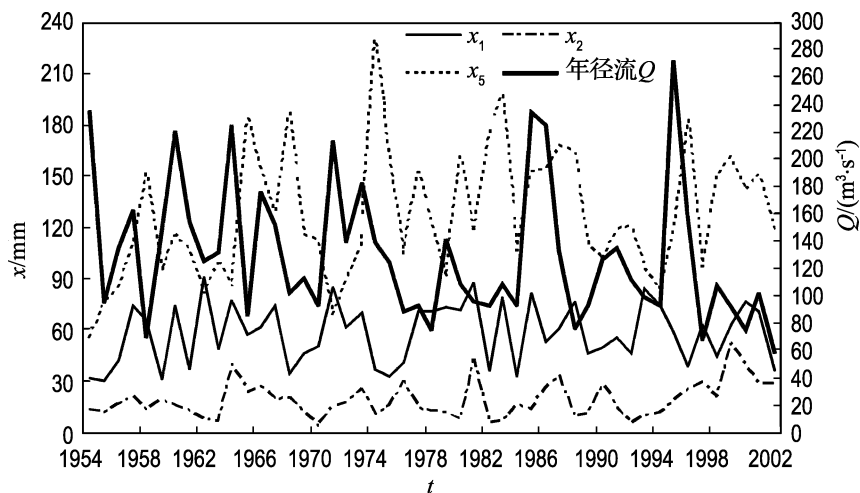


图 3 桓仁水库实测年径流量及其 3 个预报因子特征值

Fig. 3 Practical annual runoff and its three forecast factors of Huanren Reservoir

3.1.2 用复合相关系数评价与选择预报因子
预报因子的选择基于两点:其一是预报因子由少到多;其二是根据复合非线性相关系数 D 来判断是否引入一因子,即若加入一因子后对复合非线性相关系数 D 有改进,则可以引入,否则不必引入。下面设定相同的迭代误差 $\epsilon = 0.0038$,举例说明预报因子的选择方法:在因子 x_1, x_3, x_4 的基础上,即求得 $D(x_{134}) = 0.65$,确定是否选择 x_2 。将样本指标特征值规格化后的矩阵为 $\mathbf{R}$,应用公式(4)将预报量年径流量规格化后的矩阵为 $\mathbf{A}$ 。代入模糊优选神经网络模型进行计算,将模拟结果

代入式(11)得到 $D(x_{1234}) = 0.87$,进一步得 $D(x_{1234}) - D(x_{134}) = 0.22 > 0$ 。可见,引入 x_2 后 D 增加明显,对预报结果改善较大,说明预报因子 x_2 起重要作用,需要引入。

最终选择的预报因子组合为 $(x_1, x_2, x_3, x_4, x_5)$,按照文献[9]中短期预报的精度要求,若以预报误差 20% 为合格其预报合格率为 94.3%;若根据文献[9]中中长期预报的多年变幅的 10% 的许可误差要求其预报合格率为 89.2%,从图 4 中可以看出拟合趋势较好,表明选择的因子比较有效。

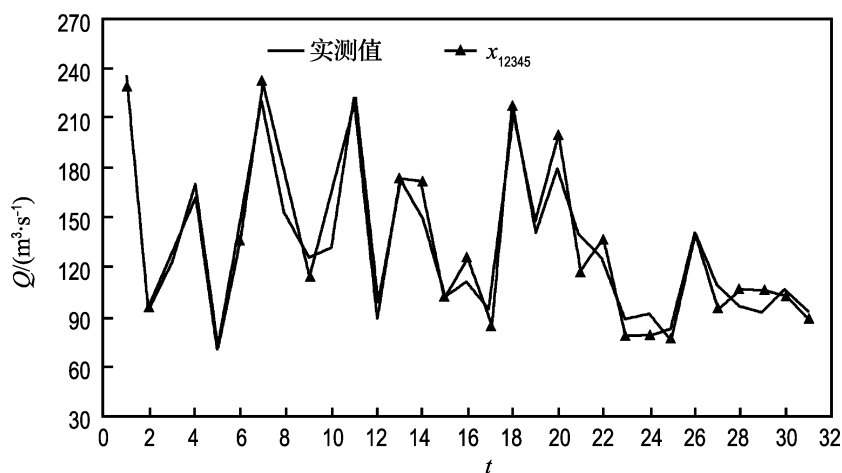


图4 多因子组合的拟合预报结果图

Fig. 4 Simulated forecast results by jointing multi-factors

由实例可见,尽管单个预报因子与预报量的相关系数都很小(7个因子依次为 0.172、0.229、0.156、-0.071、0.423、-0.097 和 0.040),但多因子的复合非线性相关系数达到 0.89,且综合预报结果较好,说明所选预报因子有效可用;若用传统的线性相关系数进行显著性检验,其显著性水平 0.05 对应的相关系数为 0.354,只有因子 5 满足要求,说明预报因子集不可用。另外还可看到,单因子相关系数较大的因子 6 并没有入选,说明不能简单按线性相关系数的大小来选择预报因子。总之,预报因子与预报量间确实存在复合非线性关系,应用复合非线性相关系数 D 来衡量预报水平是有效可行的,单因子的线性相关关系不再适用。

3.2 模型参数的识别

3.2.1 与模型结构有关的参数 模糊优选神经网络模型的主要参数有输入层的个数、隐含层神

经元个数,以及在允许迭代误差下求得的网络连接权重等。输入层即预报因子的确定前面已经阐述,隐含层节点个数的确定,Hecht-Nielsen^[11]提出为 $2N + 1$,其中 N 为输入层的节点数;Lippmann等^[12]提出为 $2M$,其中 M 为输出层的节点数,继而提出公式 $H = \log_2 T$,其中 H 为隐含层节点数, T 为输入训练模式数。在本例中, $N = 4$, $M = 1$, $T = 31$,综合上述公式得隐含层节点数分别为 9、2、5,此处取 5。

3.2.2 允许迭代误差 根据模型的综合效应系数 μ 来研究确定模型的训练参数——允许迭代误差。方法是对初步选定的预报因子组合,首先在可能的变化范围内设定多个迭代误差,分别求得与每一迭代误差相应的模型综合效应系数;然后比较分析综合效应系数的变化趋势,以确定允许迭代误差的变化范围;最后综合分析确定允许迭代

误差. 其中,只使用前 41 a 资料确定综合效应系数 μ 与允许迭代误差 ϵ .

根据已有的研究成果与初步分析,对迭代误差小于 0.001 时外推预报精度不高且计算耗时,大于 0.01 的模型拟合精度明显偏低,因此初步确定 μ 的可能变化区间为 $[0.001,0.01]$. 为此在 $[0.001,0.01]$ 区间内设定多个迭代误差,求得相应的 μ 值,其变化趋势如图 5 所示.

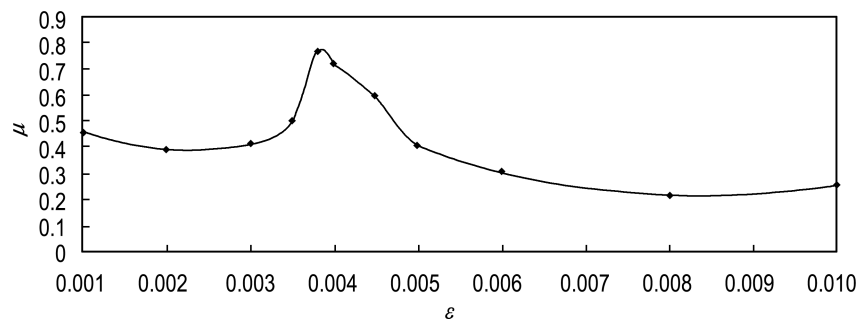


图 5 迭代误差 ϵ 与综合效应系数 μ 的相关图

Fig. 5 The relation between iterative error ϵ and μ

由图 5 可以看出,本例中综合效应系数 μ 在一定范围内存在极值,建议设定迭代误差 ϵ 在极值附近,本例在 $[0.002,0.005](\mu > 0.4)$,表 1 为选用允许迭代误差 $\epsilon = 0.0038$ 的检验样本预报结果,其 31 个训练样本的拟合复合非线性相关系数为 0.89,检验样本的复合非线性相关系数为 0.93, $\mu = 0.770$.

表 1 预报拟合结果($\epsilon=0.0038$)

Tab.1 The fitting results of test samples ($\epsilon=0.0038$)

样本序号	实测值/ ($\text{m}^3 \cdot \text{s}^{-1}$)	预报值/ ($\text{m}^3 \cdot \text{s}^{-1}$)	相对误差/%
32	234.0	225.0	3.9
33	225.0	199.0	11.3
34	133.0	144.0	8.6
35	75.0	86.2	14.9
36	92.6	107.0	16.0
37	127.0	119.0	6.6
38	135.0	132.0	2.1
39	112.0	92.9	17.1
40	98.6	111.0	12.5
41	93.2	77.8	16.5

注:平均相对误差为 10.9%;复合非线性相关系数为 0.93

以最后 8 a 作为外推预报检验样本,据前述设定的 $\mu = 0.770$ 、 $\epsilon = 0.0038$ 的模型参数,得到预报结果列于表 2. 由表 2 知,8 a 的组合预报因子与预报量间的复合非线性相关系数高达 0.91,外推预报平均相对误差只有 12.7%,表明模型具有较强的外推预报能力.

表 2 外推预报检验结果($\mu=0.770, \epsilon=0.0038$)

Tab.2 Forecast results of verifying samples
($\mu=0.770, \epsilon=0.0038$)

样本序号	实测值/ ($\text{m}^3 \cdot \text{s}^{-1}$)	预报值/ ($\text{m}^3 \cdot \text{s}^{-1}$)	相对误差/%
42	272.0	229.0	15.8
43	156.0	137.0	12.2
44	67.5	79.7	18.1
45	107.0	118.2	10.5
46	91.1	80.4	11.8
47	73.9	79.1	7.0
48	102.0	109.0	6.9
49	57.0	68.2	19.6

注:平均相对误差为 12.7%;复合非线性相关系数为 0.91

4 结 语

本文从影响年径流预报的预报因子出发,在根据当地预报经验初选预报因子的基础上,结合模糊优选神经网络预测方法,对预报因子和预报量的相关性进行探讨,在验证常用的线性相关系数不适用的基础上,提出一种复合非线性分析方法选择预报因子,结果表明这种方法是有效可行的.

在处理预报中拟合精度高而预测精度低的问题过程中,本文建立了一个综合效应系数 μ 用于反映模型的综合预报水平,在模糊优选神经网络预测模型中进行应用,根据 μ 变化趋势进行分析确定. 在应用实例中,得到对于模糊优选神经网络方法应用到年径流预报的情况,允许迭代误差 ϵ

设定在 $[0.002, 0.005]$ 为宜.

本研究提出的预报因子和预报量间相关分析的方法可供其他中长期预报模型参考应用, 综合效应系数法评价模型的预报水平、确定模型参数的方法对其他基于数据驱动的模拟模型具有借鉴意义.

参考文献:

- [1] 陈守煜. 中长期水文预报综合分析理论模式与方法[J]. 水利学报, 1997(8): 15-21
- [2] 金菊良, 魏一鸣, 丁 晶. 投影寻踪门限回归模型在年径流预测中的应用[J]. 地理科学, 2002, 22(2): 171-175
- [3] IMRIE C E, DURUCAN S, KORRE A. River flow prediction using artificial neural networks: Generalisation beyond the calibration range [J]. *Journal of Hydrology*, 2000, 233(1-4): 138-153
- [4] RIAD S, BOUCHAOU L. Rainfall-runoff model using an artificial neural network approach [J]. *Mathematical and Computer Modeling*, 2004, 40(7-8): 839-846
- [5] RAMAN H, CHANDRAMOULI V. Deriving a general operating policy for reservoirs using neural networks [J]. *Journal of Water Resources Planning and Management*, ASCE, 1996, 122(5): 342-347
- [6] 楚永伟, 蓝永超, 李向阳, 等. 黑河莺落峡站年径流长期预报模型研究[J]. 中国沙漠, 2005, 25(6): 869-873
- [7] 张利平, 王德智, 夏 军. 白山水库径流中长期预报研究与应用[J]. 水电能源科学, 2002, 20(1): 18-20
- [8] 陈守煜. 工程模糊集理论与应用[M]. 北京: 国防工业出版社, 1998: 51-54.
- [9] 中华人民共和国水利部. SL250-2000 水文情报预报规范[S]. 北京: 中国水利水电出版社, 2000: 18-19
- [10] HONIK K. Approximation capabilities of multilayer feed forward networks [J]. *Neural Networks*, 1991, 6(8): 1069-1072
- [11] HECHT-NIELSEN R. Applications of counterpropagation networks[J]. *Neural Networks*, 1988, 1(12): 131-139
- [12] LIPPMANN R P, CUNNINGHAM R K. Improving intrusion detection performance using keyword selection and neural networks[J]. *Computer Networks — The International Journal of Computer and Telecommunication Networking*, 2000, 34(4): 597-603

Research on application of fuzzy optimization neural network model to medium-term and long-term runoff forecast

LIU Yan-li, YUAN Jing-xuan, ZHOU Hui-cheng*

(School of Civil and Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: Medium-term and long-term runoff forecast is significant to reservoir operation and water resources utilization. In order to improve the lower level of medium-term and long-term runoff forecast, a fuzzy optimization neural network model is imported. A primary issue is how to choose factors since the common linear coefficient is not adaptable for the nonlinear system. Another issue is that higher level of fitting precision and lower level of forecasting precision coexist. By defining a nonlinear coefficient, the complex nonlinear relations are analyzed between forecast factors and forecast object. The effective factors are chosen according to the nonlinear coefficient. Integrative effective coefficient is defined to evaluate the effect of forecast model. It provides a solution for the problem, higher level of fitting precision and lower level of forecasting precision.

Key words: fuzzy optimization; neural network; annual runoff forecast; selection of forecast factors; integrative effective coefficient