

三种神经网络在洪水预报中应用的比较^①

侯翔¹, 汤元斌², 刘笃晋¹, 江芝蒙²

¹(四川文理学院 计算机科学系, 达州 635000)

²(四川文理学院 现代教育技术中心, 达州 635000)

摘要: 本文以四川省达州市州河流域的洪水为研究对象, 分别采用标准 BP 算法、Levenberg-Marquart 算法和遗传算法来建立洪水预报模型, 并对预报结果进行了分析和比较。结果表明: 三种算法之中, 遗传算法所建立的模型的收敛速度最快, 预测结果精度最高, 能够有效避免网络陷入局部极小点。

关键词: 神经网络; BP 算法; Levenberg-Marquart 算法; 遗传算法; 洪水预报

Comparative Study on the Applications of Three Neural Networks to Flood Forecasting

HOU Xiang¹, TANG Yuan-Bin², LIU Du-Jin¹, JIANG Zhi-Meng²

¹(Department of Computer Science, Sichuan University of Arts and Science, Dazhou 635000, China)

²(Modern Educational Technology Center, Sichuan University of Arts and Science, Dazhou 635000, China)

Abstract: By establishing flood forecasting models based on the standard BP algorithm, the Levenberg-Marquart algorithm and the genetic algorithm respectively, the paper studies the floods of the Zhouhe River in Dazhou city of Sichuan province, and analyzes and compares the outcomes of these three models. It shows that the model based on the genetic algorithm has the fastest convergence rate and highest accuracy in flood forecasting and can effectively prevent the network from getting into local minimum point.

Key words: neural network; BP algorithm; Levenberg-Marquart algorithm; genetic algorithm; flood forecasting

1 引言

洪水预报是水文资源学科中的重要分支, 依据水文河观测站监测暴雨或上游的洪水, 经过水情信息的采集、传输、处理等复杂的科学运算, 最终准确预报在流域的出口处或下游的水文观测站实际洪水的过程, 帮助水利部门有效地管理水资源, 降低洪水带来的风险, 保障人民群众的生命财产安全^[1]。洪水预报确保水库安全, 保障城市用水, 减轻洪水带来的一系列灾害, 对整个流域的防洪安全有着非常重要的意义。但是在实际洪水预报中由于受到多方面的影响, 如地理位置复杂、突发事件和人们活动等, 因此一直以来流域洪水预报是水文预报工作中的重点难点问题^[2]。

一般说来洪水预报有三种不同的方法: 经验法, 统计法(平均值法、流量等级法)以及模型方法。经验方法是人们早期根据洪水发生特点和发生时间来预报, 这种方

法因人而异, 不稳定差别比较大, 效果不好^[3,4]。统计方法就是统计近些年流域发生洪水的时间, 流量等等因素进行分析, 最后进行分析预报。而模型方法中, 以美国为代表的研究洪水预报模型有斯坦福(Stanford)模型、萨克拉门托模型, 日本也提出了坦克(Tank)模型, 以及中国研究人员也提出的新安江模型^[5-7]。美国研究人员提出根据递归RAR自适应更新模型, 比传统的时间系列预报模型精度更高, 另外由于是自适应的, 参数可以调整, 可以实时更新预报模型调整预报误差等方法。

这些传统洪水预报方法都需要大量的径流、降雨等资料, 其影响预报结果的水文因子众多, 与预报对象之间的关系复杂, 建立的过程通常需要较久的时间, 而且在处理输入数据时主要使用了线性关系, 缺乏自适应性, 模型更新困难, 预报精度时好时坏, 不是十分稳定。事实说明, 水文系统中很多变量间的复杂非

^① 基金项目:四川省教育厅 2011 年面上项目(11ZB139);达州市 2011 年科技攻关项目(JCY1117)

收稿时间:2013-05-08;收到修改稿时间:2013-05-30

线性关系远非简单的代数方程所能有效表达。

近些年,随着科学计算、智能算法、云计算等方法的发展,人们不断的丰富洪水预报模型。人工神经网络具有信息分布存贮和并行处理等特点及自组织、自学习的功能,利用其极强的非线性函数逼近能力,能对难以精确描述的复杂非线性对象进行建模,从而实现了对复杂非线性对象进行有效的模拟。因此,神经网络在洪水预报上得到了十分广泛的应用,只要给出以前的水文资料,神经网络就可通过自身的学习来完成洪水的预报工作^[8]。

BP网络是目前应用最广泛的神经网络模型之一,但是标准的BP网络收敛速度慢、易陷入局部极小值、难确定隐含层数和隐节点数等^[9],因此,人们尝试采用各种算法对BP网络进行了优化,例如:可变学习率算法、贝叶斯正则化法、蚁群算法等。本文以四川省达州市州河流域的洪水为研究对象,选取其中最具有代表性的2种算法:遗传算法和Levenberg-Marquart算法(L-M算法),与标准BP算法进行了分析和比较。

2 研究方法

(1) 标准BP算法。BP算法通常为3层结构,也可以是多层结构,由输入层、输出层和1个隐含层或者多个隐含层组成。其学习过程分为信号正向传播和误差信号反向传播2个阶段。BP算法默认采用了梯度下降法,但是存在线性收敛速度慢、容易陷入局部极小点等缺陷。

(2) L-M算法。针对标准BP算法线性收敛速度慢,结合了BP神经网络的最速下降法和Gauss-Newton法,因此,具有Gauss-Newton法的局部收敛能力,又具备了最速下降法的全局特性。通过自适应调整阻尼因子来达到收敛,提高了收敛速度,训练速度要比标准BP算法快^[10]。

(3) 遗传算法(Genetic Algorithm)。针对标准的BP网络随机生成初始权值,预报结果具有随机性,没办法判断是否最优解的去电,结合遗传算法能够同时搜索函数曲线中多个个体,全局搜索能力极强,不容易陷入局部最小点的特点,是一种通过模拟适者生存,优胜劣汰的规律,在全局上搜索最优解的算法。它们摆脱了对函数的依赖性,通过对种群个体进行遗传操作,从而实现种群内个体结构重组的迭代过程。在这个过程中,种群个体一代代的得以优化并逐渐接近最佳解。

3 应用实例

针对洪水预报的问题,只要选定了预报对象和预报因子,就可以建立模型,对预报对象进行预报。就本次基于三种不同算法的洪水预报模型而言,将流域下游出口流量作为预报对象。

3.1 流域概况

州河,如图1所示,始于宣汉县江口,经宣汉城南门、西北、东林、洋烈至千丘旁入达县境。由东北向西南,经达县罗江乡红梁村曹家湾进入达县境内,穿过达州市和达县的罗江、河市、渡市三个区的7个乡镇,于木头乡的大河咀出境,流入渠县的农乐、汇东、汇南等乡,在三汇镇与巴河相汇,干流由前河、中河和后河汇聚而成。州河流域面积8849 km²,河长108km。多年平均流量190m³/s,据历史洪水调查最大流量13700 m³/s(1902年),多年平均径流总量60.1亿m³。河宽一般为200~300米^[11]。

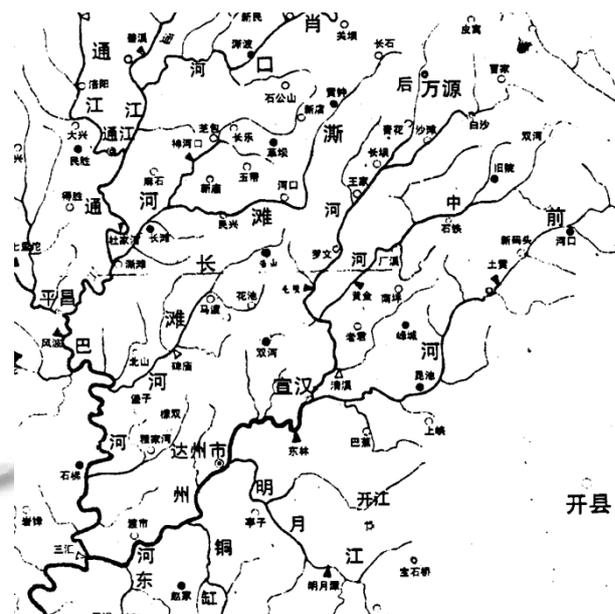


图1 州河水系图

3.2 神经网络模型的建立

通常上游洪水过程与下游洪水过程具有很好的相关性,因此在做下游洪水流量预报时,上游过去洪水流量过程应该是重要的影响因子。同时,在利用当地的时间序列建模,也能获得较好的预报结果,故在洪水流量预报时,当地过去时刻的流量过程也是重要的影响因子。因此,在做下游洪水流量预报时,上游和当地过去洪水流量过程都是重要的影响因子。

在州河流域范围内,根据《水文情报预报规范》对资料的要求,按照易获取、资料连续、代表性强等原则,筛选出上游水文站土黄(前河)、黄金(中河)、毛坝(后河)和下游水文站东林(干流)共计 4 个站点作为该流域的代表站,采用 2001 年~2006 年间 4 个水文站的部分洪水的流量作为训练样本,选择 2007 年的流量作为测试样本.以土黄站、毛坝站、黄金站的洪水流量 $Q_1(t)$ 、 $Q_2(t)$ 、 $Q_3(t)$ 为模型输入,预报东林站洪水流量 $Q(t+\Delta t)$,模型的数学公式为:

$$Q(t+\Delta t) = f[Q_1(t), Q_2(t), Q_3(t)] + \varepsilon(t+\Delta t) \quad (1)$$

其中,式中 Q 、 Q_1 、 Q_2 、 Q_3 分别为东林站、土黄站、毛坝站、黄金站的洪水流量 m^3/s .

(1) 本研究的对象是洪水,它具有非线性关系的特征,因此在 BP 网络建立时,传递函数采用 Sigmoid 函数.但由于 Sigmoid 函数的特性,为了加快 BP 网络的收敛性,有必要对所有的输入值加以归一化.如果不对数据进行归一化处理,直接输入较大的流量值到神经网络中,将可能导致浮点误差及训练时间增长^[12].归一化运算的公式为:

$$y = \frac{(p - p_{\min})}{(p_{\max} - p_{\min})} \quad (2)$$

其中, p 为当前学习样本; p_{\max} 为学习样本中的最大值; p_{\min} 为学习样本中的最小值; y 为训练归一化后的数据.

(2) 模型率定.在模型参数的优化识别中,目标函数的选取直接影响优化计算的结果和优化计算的速度.本文选用率定期内洪水径流总量、流量过程误差最小、合格率最高等多目标函数来优选模型参数.

研究表明,单隐层的 BP 神经网络能以任意评价参数设定的精度逼近函数,因此选择单隐含层前馈网络,即三层 BP 网络,并通过试算,确定网络结构为 3-9-1.三种 BP 算法的训练参数优选为:预报精度 $goal=0.0001$,学习速率 $lr=0.001$,训练次数 $epochs=20000$,其余参数取默认值.遗传算法在运行过程中的重要参数设置为:种群规模为 200,迭代次数为 2000,选择概率为 0.1,交叉概率为 0.3,变异概率为 0.1.

3.3 模型评估指标

采用《水文情报预报规范》(SL250-2000)中的规定的合格率和确定性系数作为模型评估指标^[13].一次洪水预报的误差小于许可误差即为合格预报.合格率 QR 表示经过多次预报后总的精度水平,确定性系数

DC 表示洪水预报过程与实测过程之间的拟合程度,模型的精度按照仿真结果的确定性系数或合格率可被分为三个不同的等级,见表 1 所示.公式为:

$$QR = \frac{n}{m} \times 100\% \quad (3)$$

$$DC = 1 - \frac{\sum_{i=1}^n [y_c(i) - y_0(i)]^2}{\sum_{i=1}^n [y_0(i) - \bar{y}_0]^2} \quad (4)$$

式中, m 是预报总次数; n 是合格预报次数. $y_0(i)$ 为实测值; $y_c(i)$ 预报值; \bar{y}_0 为实测值的平均值.

表 1 预报项目精度等级表

精度等级	甲	乙	丙
合格率	$QR \geq 85.0$	$85.0 > QR \geq 70.0$ 0	$70.0 > QR \geq 60.0$ 0.0
确定性系数	$DC \geq 0.90$	$0.90 > DC \geq 0.7$ 0	$0.70 > DC \geq 0.5$ 50

3.4 预报结果的分析 and 比较

三种 BP 算法的预测流量与实测流量的比较,如图 2 所示,两种优化 BP 算法的预测流量与实测流量十分吻合,预报效果比标准 BP 算法更好.

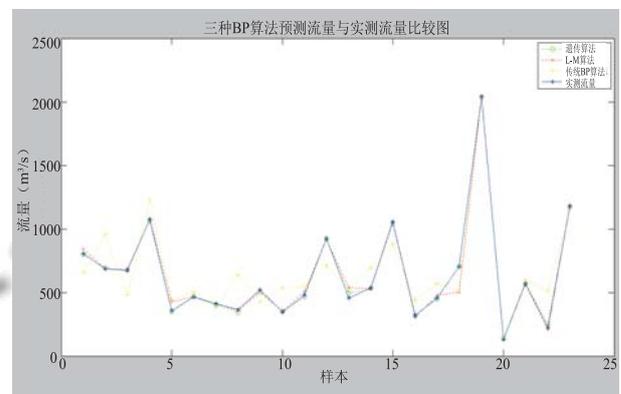


图 2 三种 BP 算法预测流量与实测流量比较

三种 BP 算法的具体项目比较见表 2. 由表可知,遗传算法和 L-M 算法的 DC 和 DR 均优于标准 BP 算法,预报项目精度等级为甲级,而标准 BP 算法其预报项目精度等级仅仅为乙级.遗传算法在执行 254 步、学习 10 秒后,便能得到预期精度,收敛速度在三种算法中最快; L-M 算法则需要训练 915 步才能够收敛于预报精度,学习时间为 12 秒,比遗传算法的收敛速度略慢;普通 BP 算法则陷入局部极小值,执行到预设的最大执行步骤 20000 步时,程序自动终止,未能得到

满意结果。由上述分析可知,两种优化 BP 算法的预报精度和收敛速度都较标准 BP 算法有较大的提高。

表 2 三种 BP 算法的比较

比较项目	标准 BP 算法	L-M 算法	遗传算法
训练次数	20000	915	254
学习时间	0:01:47	0:00:12	0:00:10
预报精度	0.0181	0.00105	0.001008
合格率	78.26%	100%	100%
确定性系数	0.7483	0.9958	0.9985
绝对误差和	1836.60	77.7771	3.0200
相对误差和	0.0106	0.00291	0.00097

4 结论

人工神经网络在洪水预报中应用前景广阔,但是从众多的神经网络中选取最适合建立洪水预报模型的算法很困难。本文以四川省达州市州河流域的洪水为研究对象,采用基于 L-M 算法和遗传算法优化的 BP 神经网络建立洪水预报模型,并与普通 BP 算法的预报结果进行了比较。从预报结果分析,两种优化后的 BP 算法收敛速度和预测精度均比标准 BP 算法有较大的提高,能够有效的避免 BP 网络陷入局部极小点,而遗传算法的收敛速度在三种算法之中是最快的,预测精度也是最高的,因此,遗传算法是最适合于建立洪水预报模型。

参考文献

1 Coulibaly P, Baldwin CK. Nonstationary hydrological time

(上接第 18 页)

- 工程,2003,(3):91-94.
- 2 Premachandra IM. An approximation of the activity duration in PERT. *Computers & Operations Research*, 2001, 28(5): 443-452.
- 3 钟登华,刘奎建,杨晓刚.施工进度计划柔性网络仿真的不确定性研究. *系统工程理论与实践*,2005(2):107-112.
- 4 王仁超,欧阳斌,王琳,褚春超.工程项目计划“关键性”问题拓展研究. *系统工程与电子技术*,2004-7,26(7):914-923.
- 5 王祖和,亓霞.多资源均衡的权重优选法. *管理工程学报*, 2002,16(3):91-93.
- 6 李平,顾新一.PERT 网络工程风险计算方法的研究. *统计与*

series forecasting using nonlinear dynamic methods. *J. Hydrol*, 2005, 307: 164-174.

- 7 赵玉庆.模糊神经网络及遗传算法在实时洪水预报中的应用研究[硕士学位论文].大连:大连理工大学,2012.
- 8 范睿.基于遗传算法的神经网络洪水预报研究与应用[硕士学位论文].哈尔滨:哈尔滨工程大学,2005.
- 9 金保明.BP 神经网络在闽江十里庵流量预测中的应用. *水电能源科学*,2010,28(9):12-14.
- 10 文雨松,李整,邹希云等.基于桥墩历史水痕与气象预报的中小桥水害预测. *铁道学报*,2010,32(2):141-144.
- 11 王艳君,金生.河网水情预测的三种 BP 神经网络方法比较. *水电能源科学*,2010,28(2):19-21.
- 12 朱全银,戴峻峰.洪泽湖水位预测模型的研究. *计算机仿真*,2009,26(4):113-115.
- 13 阎平凡,张长水.人工神经网络与模拟进化计算.北京:清华大学出版社.2000:26-27.
- 14 隋彩虹,徐宗学.GABP 模型在渭河下游洪水预报中的应用. *人民黄河*,2007,(4).
- 15 殷峻逞,陈守煌,邱菊.基于遗传与 BP 混合算法神经网络预测模型及应用. *大连理工大学学报*,2002,42(5):594-595.
- 16 百度百科.渠江. <http://baike.baidu.com/view/746730.htm>, 2012,9.
- 17 飞思科技产品研发中心.神经网络理论与 MATLAB7 实现.北京:电子工业出版社.2005.
- 18 中华人民共和国水利部.SL250-2000 水文情报预报规范.北京:中国水利水电出版社.2000.
- 19 决策,2004(5):16-17.
- 20 王伟,蔡晨.在两资源约束项目环境中的关键链管理. *中国管理科学*,2003-10,11:4-8.
- 21 马国丰,屠梅曾,史占中,吴伟.基于关键链技术的项目进度管理系统设计与实现. *上海交通大学学报*,2004(3):377-381.
- 22 陈魁.应用概率统计.北京:清华大学出版社.2000.
- 23 曹成军,别朝红,王锡凡.蒙特卡洛法全周期抽样的研究. *西安交通大学学报*,2002,(4):344-352.
- 24 唐爱国,王如龙.基于贝叶斯网络的软件项目风险评估模型. *计算机工程*,2008,(22):91-93.