

【水文·泥沙】

# BP神经网络在大理河年径流预测中的应用

赵丽娜, 宋松柏

(西北农林科技大学 水利与建筑工程学院, 陕西 杨凌 712100)

**摘 要:**以大理河为例, 根据其 1964—2001 年年径流序列的长期变化特征, 利用 MATLAB 工具箱提供的神经网络设计、训练及仿真函数建立了年径流预测的 BP 神经网络预测模型。结果表明: BP 神经网络预测值检验的合格率为 100%, 预报精度明显高于自回归模型。

**关 键 词:** MATLAB 工具箱; BP 神经网络; 年径流; 大理河

**中图分类号:** P338.6    **文献标识码:** A    doi: 10.3969/j.issn.1000-1379.2010.08.013

大理河位于陕西北部, 干流全长 170 km, 流域面积为 3 906 km<sup>2</sup>, 是无定河最大的支流, 该流域大部分属黄土丘陵沟壑区, 地形破碎, 植被稀疏, 水土流失严重。预测其年径流量对于合理开发和利用径流资源, 更好地制定区域社会经济发展规划具有十分重要的意义, 笔者应用 BP 神经网络对此进行了分析。

## 1 理论基础

### 1.1 BP神经网络的建模原理

目前, 在人工神经网络<sup>[1-5]</sup>的实际应用中, 绝大部分神经网络模型采用 BP (Back Propagation) 网络及其变换形式, 它是前向网络的核心部分。BP 网络是一种有隐含层的多层前馈模型, 能系统地解决多层网络中隐含单元连接权的学习问题。BP 学习算法的基本原理是梯度最速下降法, 其中心思想是调整权值使网络总误差最小。网络学习过程是一种误差一边向后传播一边修正权系数的过程。

神经元是神经网络最基本的组成部分, 一般来说, 一个有  $n$  个输入的神经元模型结构见图 1。其中:  $P_i$  为输入;  $W_{ji}$  为权系数;  $b$  为阈值;  $F$  为转移函数;  $T$  为神经元输出。传递函数可以是任意可微的函数, 常用的有 Sigmoid 型和 purelin 型。

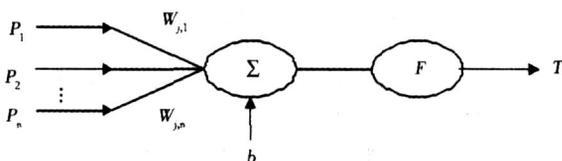


图 1 神经元模型

### 1.2 基于 BP 网络的时间序列预测的具体步骤<sup>[6-9]</sup>

(1) 网络拓扑结构的设计。多个神经元相互连接组合构成一个神经网络体系, 可对复杂的信息进行有效的识别处理。这种神经元的连接方式称为人工神经网络模型的拓扑结构。网络结构的设计是指确定网络的隐含层数、隐含层节点数和输入输出节点数。Hecht-Nielsen 于 1987 年指出: 1 个 3 层 BP 网络可以满足一般函数的拟合逼近问题, 因此网络的层数选择 3 层。输入层的节点数是由其影响因子的个数决定的。采用试

验法确定隐含层的节点数, 即不断增加隐含层的节点数进行网络训练和测试, 比较不同方案的训练和测试结果, 选取合适的隐含层节点数。由于本研究限于一维时间序列, 因此只需要一个输出层节点。

(2) 输入输出数据的预处理。由于网络的转换函数 (Sigmoid 函数) 在  $[0, 1]$  区间的变化梯度较大, 因此一般网络训练时均把样本数据变换到这一区间。对输入输出向量作如下处理:

$$P_i = \frac{(P_i - P_{\min}) \times 0.996}{P_{\max} - P_{\min}} + 0.002 \quad (1)$$

$$T = \frac{(T - T_{\min}) \times 0.996}{T_{\max} - T_{\min}} + 0.002 \quad (2)$$

式中:  $P_i$  和  $T$  为原始的输入、输出样本;  $P_i$  和  $T$  为处理后的输入输出样本;  $P_{\min}$  和  $P_{\max}$  分别为第  $i$  个影响因子在样本中的最小值和最大值;  $T_{\min}$  和  $T_{\max}$  为原始输出样本的最小值和最大值。

(3) 网络训练。给网络赋予一个初始权值, 用训练权值和阈值按照 BP 算法的步骤, 对网络进行反复训练, 直到网络收敛于一定的标准。否则, 可以改变网络的初始权值甚至网络的拓扑结构, 直到达到满意的误差为止。

(4) 网络泛化能力的检验。当网络训练完成后, 训练好的网络权值和阈值固定保存在知识库中。这时整个网络就是一个时序预测模型, 用检验样本对该网络模型的泛化能力进行检验, 评价预测精度最常用的指标是 MAPE (Mean Absolute Percentage Error):

$$MAPE = (1/N) \sum_{k=1}^N |(P_k - \hat{P}_k) / P_k| \times 100\% \quad (3)$$

式中:  $P_k$  为检验样本实测值;  $\hat{P}_k$  为预测值。

收稿日期: 2009-05-31

作者简介: 国家自然科学基金资助项目 (50879070, 50579065); 西北农林科技大学青年学术骨干支持计划和优秀博士论文基金资助项目 (Z2007)。

作者简介: 赵丽娜 (1983—), 女, 内蒙古包头人, 硕士研究生, 研究方向为水文水资源。

通讯作者: 宋松柏 (1965—), 男, 教授。

E-mail: ssb6533@yahoo.com.cn

## 2 MATLAB下神经网络的实现

在 MATLAB设计环境中设计 BP网络<sup>[10]</sup>时,只要已知输入向量 P、各层的神经元个数及各层神经元的转换函数,就可以利用函数 intff对 BP网络进行初始化,其初始化语句(n为隐含层的节点数)为

```
[w1, b1, w2, b2] = intff(P, n, 'tansig', T, 'purelin')
```

神经网络工具箱 trainbp, trainpx, trainm 均可用来对 BP网络进行训练。笔者采用 trainbp函数进行训练。前馈网络由一系列网络层组成,每一层都是从前一层得到输入数据, simuff函数用来模拟 3层前馈网络,在 MATLAB中的调用函数语句(Q为输入数据, a为模拟数据)为

```
a = simuff(Q, w1, b1, 'tansig', w2, b2, 'purelin')
```

## 3 实例应用

### 3.1 基本资料

以大理河绥德站共 38年(1964—2001年)的年径流资料与年降水量、年平均相对湿度为研究数据建立模型。以 1964—1998年的实测资料为样本训练网络,以 1999—2001年的实测年径流资料为检验样本对模型加以检验。

### 3.2 数据处理

根据式(1)对影响因子年降水量、年平均相对湿度样本数据作归一化处理,同时为了防止部分神经元达到过饱和状态,根据式(2)对输出样本作同样的处理。

### 3.3 MATLAB中BP神经网络结构的实现

据资料分析可知,影响该地区年径流量的主要因素有当年年降水量和当年年平均相对湿度。另外,根据分析实测年径流量的相关关系,认为每年的径流量与前两年的年径流量相关性较强,而与其他年的相关性较弱,故选取前两年的年径流数据作为输入样本,当年数据作为输出样本。因此,设定输入层的节点数为 4。经过反复试验,隐含层的节点数确定为 53。隐含层利用 Sigmoid型函数,输出层利用 purelin型函数。

通过调用函数 intff对 BP网络进行初始化,该 BP网络的初始化语句为

```
>> [w1, b1, w2, b2] = intff(P, 53, 'logsig', T, 'purelin')
```

通过调用函数 trainbp实现网络训练,训练函数的参数设置如下:

```
>> df = 20;
>> lr = 0.001;
>> me = 10000;
>> eg = 0.1;
>> tp = [df me eg lr];
>> [w1, b1, w2, b2, ep, tr] = trainbp(w1, b1, 'logsig', w2, b2, 'purelin', P, T, tp)
```

由运行结果可以看出,网络误差达到收敛于 0.1的要求。

通过调用函数 sim实现网络的仿真,该仿真语句为

```
>> a = simuff(P, w1, b1, 'logsig', w2, b2, 'purelin')
```

调用 plot函数,绘制神经网络拟合图,其程序如下:

```
>> B = 1966:1:1998;
```

```
>> plot(B, a, B, T)
```

运行结果见图 2。

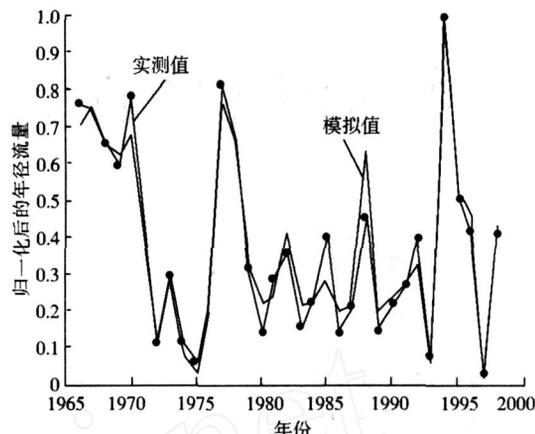


图 2 年径流量神经网络拟合

当网络训练好以后,将待预测的数据输入网络进行预测,其输入语句(Q为预留的测试样本, O为预测值)为

```
>> O = simuff(Q, w1, b1, 'logsig', w2, b2, 'purelin')
```

得到的预测结果见表 1,同时利用随机水文学的基本原理<sup>[11-13]</sup>建立了自回归 AR(1)模型对年径流量进行预测,也将预测结果列于表 1中。

表 1 年径流量数据序列的预测结果及误差

年份	日均流量 合计值/ (m <sup>3</sup> ·s <sup>-1</sup> )	BP网络		AR(1)	
		预测值/ (m <sup>3</sup> ·s <sup>-1</sup> )	相对 误差/%	预测值/ (m <sup>3</sup> ·s <sup>-1</sup> )	相对 误差/%
1999	877.074	805.3635	-8.2	702.0382	-20.0
2000	895.208	893.7705	-0.2	855.5941	-4.4
2001	1847.005	2093.3446	13.3	1586.3698	14.1

拟建的 BP神经网络预测值检验的 MAPE为 7.23%,检验合格率为 100%; AR(1)模型的预测值检验的 MAPE为 17.53%,检验合格率为 100%。BP神经网络在预报精度上明显高于自回归模型。

## 参考文献:

- [1] 张少文,张学成,德格吉玛,等.黄河上游天然年径流长期变化趋势预测[J].人民黄河,2007,29(1):27-30
- [2] 顾海燕,徐文科,于雷.基于BP神经网络的河川年径流量预测[J].东北林业大学学报,2007(10):83-85
- [3] 钟登华,王仁超,皮钧.水文预报时间序列神经网络模型[J].水利学报,1995(2):69-75
- [4] 庞博,郭生练,熊立华,等.改进的人工神经网络水文预报模型及应用[J].武汉大学学报:工学版,2007(1):33-41
- [5] 缪益平,邓俊.基于BP人工神经网络的枯水径流预报方案研究[J].水文,2008(3):33-37
- [6] 周振民,刘荻.基于MATLAB的人工神经网络用水量预测模型[J].中国农村水利水电,2007(4):45-49
- [7] 王玲,黄国如.基于径流分类的日径流量预测神经网络模型[J].灌溉排水,2002(4):45-48
- [8] 马细露.神经网络处理多变量关系的应用分析[J].人民黄河,2005,27(12):36-39
- [9] 张倩倩,姜铁兵,康玲,等.基于人工神经网络的日径流预测[J].水电自动化与大坝检测,2002(4):65-67

(下转第 33 页)

沉速大于单颗粒泥沙的沉速;随着含沙量的继续增大,絮团与絮团之间形成一个连续的空间结构网,出现一定的刚性,这时泥沙的沉速大幅降低;空间结构网刚形成时仅由极细沙组成,它们与清水组成均质的浆液,以极其缓慢的速度下沉,粗颗粒泥沙虽受絮凝结构的影响而减小了沉速,但依然保持分散体系的性质自由下沉,在沉降过程中存在着粗、细泥沙的分选;含沙量继续增大以后,越来越多的较粗颗粒泥沙也成为絮凝结构的一部分,自由沉降部分所占的比例越来越小;含沙量继续增加达到某个临界值时,全部泥沙均参与组成均质浆液这一过程,这样便不存在粗细泥沙的分选。张浩等采用中径为 0.035 mm 的混合沙进行的试验表明,当含沙量达到 800 kg/m<sup>3</sup> 时,泥沙无论粗细均作为一个整体下沉,已看不到分选现象。

### 3.2 河床层理淤积机理探讨

费祥俊<sup>[11]</sup>的研究表明:细颗粒淤积物在沉积过程中受絮凝作用影响会连接成絮团,絮团与絮团会连接成集合体,集合体还会搭接而形成网架。絮凝的新沉积物是一个蜂窝状的结构,含水量很高,密度很低,这样的淤积物具有很低的抗剪强度或黏结力。在自重或其他外力的作用下,最脆弱的集合体与集合体之间的连接将首先破裂,并改变沉积物结构达到较为密实的平衡状态,这样的淤积物具有较大的密度和黏结力;进一步增加压力,将使絮团之间的连接破裂,絮团集合体的形式不复存在,许许多多 7 的絮团重叠排列成层;再进一步增加压力,絮团将发生变形,絮团间孔隙消失,淤积物成为颗粒密集排列的均匀结构。显然,同一细颗粒淤积物在不同状态下所具有的密度、黏结力、抗剪强度也是不同的,甚至有极大的差别。相应地,它们抵抗冲刷的能力也大不一样。

对于黄河这类多沙河流,河道冲淤主要取决于汛期洪水,而不同的洪水来源和组成具有不同的冲淤特性。黄河北干流和渭河细泥沙来源区的洪水历时长,极细沙含量较高。相反,粗泥沙支流来的洪水峰型尖瘦,历时短、涨落急剧、含沙量高、泥沙粒径大,相应的细泥沙含量较低。这两种高含沙洪水出现的几率都较高,给极细泥沙形成絮凝沉降创造了条件。但是,来自粗泥沙来源区的洪水因极细沙含量小,形成絮凝沉降的几率小,即使形成絮凝沉降,絮团尺寸也不会很大,在沉淀、压密的过程中不会出现大面积的胶泥层,只会出现小型的胶泥块。相应地,如果洪水来自于细泥沙来源区,那么高含沙水流密度的增大必将出现较大尺寸絮凝的沉降,因而形成成片的、密实度极大的胶泥层沉积。

另外,河床的层理淤积还与洪水漫滩状况有关。洪水漫滩后,粗沙快速沉淀形成粗沙层,而细沙则缓慢沉积,覆盖于上面形成细沙层,而且在河势两弯顶之间的中间部位厚度较大,上

首细泥沙的沉积厚度一般较小,下首细颗粒泥沙由于沿程运移,因此落淤厚度也会有所减小,也就形成了人们常见和常说的“透镜体淤积状态”。

## 4 结 语

(1)泥沙成层分布是在黄河这种多沙河流上普遍存在的现象,粗沙、粉沙、胶泥层分布不一,厚度不同,黄河下游俗称“透镜体淤积”。

(2)河床泥沙成层分布与河段的水沙条件、边界条件密切相关。水量年际变化大,水沙异源,且发生高含沙洪水的支流相对比较,各支流泥沙的粗细不同等为泥沙分层淤积的主要前提,河道边界条件进一步促进了泥沙分层淤积的形成。

(3)粗泥沙来源区的洪水因极细沙含量小,形成絮凝沉降的几率小,在沉淀、压密的过程中不会出现大面积的胶泥层;来自于细泥沙来源区的洪水,随着高含沙水流密度的增大,必将出现较大尺寸絮凝的沉降,形成成片的、密实度极大的胶泥层。在不同来源区高含沙水流交替发生的过程中,粗泥沙沉积形成通常的河床淤积物,包围或覆盖在细泥沙形成的胶泥层(块)上。

## 参考文献:

- [1] 江恩惠,曹永涛,张清.黄河高含沙洪水“揭河底”冲刷研究现状[J].人民黄河,2004,26(7):6-11.
- [2] 张庆河,王殿志,吴永胜,等.黏性泥沙絮凝现象研究述评(1):絮凝机理与絮团特性[J].海洋通报,2001,20(6):80-90.
- [3] 钱宁,万兆惠.泥沙运动力学[M].北京:科学出版社,2003.
- [4] Mehta A J, Lee S C. Problems in linking the threshold condition for the transport of cohesionless and cohesive sediment grain [J]. J. Coastal Res, 1993(1): 170-177.
- [5] 张德茹,梁志勇.不均匀细颗粒泥沙粒径对絮凝的影响试验研究[J].水利水运科学,1994(1):11-17.
- [6] 张志忠.长江口细颗粒泥沙基本特性研究[J].泥沙研究,1996(1):67-72.
- [7] 夏震寰,宋根培.离散颗粒和絮凝体相结合的沉降特性[C].第二次河流泥沙国际学术讨论会论文集.北京:水利电力出版社,1983:253-262.
- [8] 钱意颖,杨文海,赵文林.高含沙水流的基本特性[C].河流泥沙国际学术讨论会论文集(第 卷).北京:光华出版社,1981:175-184.
- [9] 张浩,任增海.高含沙水流沉降规律和阻力特性[C].河流泥沙国际学术讨论会论文集(第 卷).北京:光华出版社,1981:185-194.
- [10] 黄建维,孙献清.黏性泥沙在流动盐水中沉降特性的试验研究[C].第二次河流泥沙国际学术讨论会论文集.北京:水利电力出版社,1983:274-284.
- [11] 费祥俊.浆体与粒状物料输送水力学[M].北京:清华大学出版社,1994.

【责任编辑 翟成亮】

(上接第 31 页)

- [10] 刘叔军,盖晓华,樊京,等. MATLAB 7.0 控制系统应用与实例[M].北京:机械工业出版社,2005.
- [11] 冯国章.非相依年径流序列的随机模拟[J].西北水工程,1993(3):9-16.
- [12] 王文圣,金菊良,李跃清,等.水文水资源随机模拟技术[M].成都:四川大

学出版社,2007.

- [13] 马贵安.黄河陕县站年径流序列的随机模拟——连续枯水段统计特性研究[J].人民黄河,1991,13(2):20-23.

【责任编辑 翟成亮】