

BP 神经网络方法在灌区需水预测中的应用

黄曦妮¹, 查思慧²

(1. 江西省水利投资集团有限公司, 江西 南昌 330006;

2. 河海大学水文与水资源学院, 江苏 南京 210098)

摘 要: 以江西省廖坊水利枢纽工程灌区为例, 通过查阅年鉴、现场调查和专家咨询确定影响需水量的主要因子, 根据确定的主要影响因子构造 BP 神经网络的输入样本, 采用 Matlab 软件平台编程, 建立灌区需水模型。利用建立的模型预测 2025 年灌区的需水量, 并与其他方法的预测成果进行比较。结果表明, BP 神经网络方法在廖坊水利枢纽工程灌区需水量预测的应用上是成功的。

关键词: BP 神经网络; 需水预测; 灌区; Matlab

中图分类号: TV213

文献标识码: A

文章编号: 1004-4701(2017)02-0088-05

0 引言

灌区需水量预测对合理制定灌区水资源调度计划和调配方案具有重要的意义。国内外研究学者在需水预测方面做了许多的工作, 主要采用的研究方法有指标分析法、回归分析法、灰色预测法、支持向量机法、神经网络法等^[1-8]。BP(Back-Propagation)神经网络作为最成熟的神经网络结构之一, 凭借拥有很强的自组织、自学习、自适应能力以及算法简单易实现的特点, 成为许多领域在实际应用中的首选结构, 尤其是在预测和分类上应用成效突出。本文以江西省廖坊水利枢纽工程灌区需水量为研究对象, 在 Matlab 软件中构建 BP 神经网络需水预测模型, 对灌区需水量进行需水预测, 用检验样本对 BP 神经网络需水模型预测进行检验。结果表明, BP 神经网络模拟需水量的精度较高, 其预测结果可为灌区水资源规划调度提供科学依据。

1 BP 神经网络原理

1.1 BP 神经网络的结构

BP 神经网络是一种多层前向型神经网络^[6], 由输入层、隐含层、输出层组成, 隐含层的数目不定, 应根据实际情况来选择。研究证明, 有一个隐含层的 BP 网络

可以逼近任何一个闭区间内的连续函数^[9], 本文取一个隐含层, 采用 3 层前馈 BP 神经网络, 其组成结构见图 1。各层的神经元数分别为 n_1, n_2, n_3 构成所谓的 $n_1 - n_2 - n_3$ 结构。其中隐含层神经元个数的确定方法一般采用经验公式法, 根据 Kolmogorov 定理^[10]可知, 针对只有 3 层 BP 神经网络隐含层神经元的数量为输入层神经元数量的两倍加 1, 即组成 $n_1 - (2n_1 + 1) - n_3$ 的网络结构。

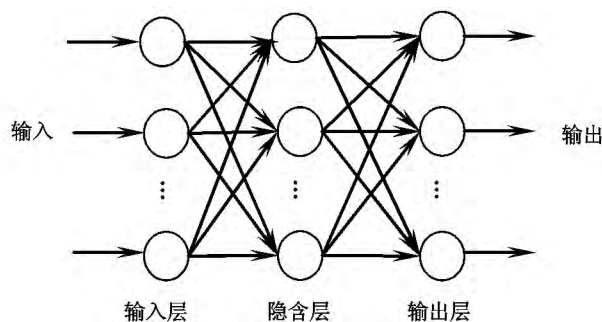


图 1 BP 神经网络 3 层结构示意图

1.2 BP 神经网络的训练算法

BP 网络的训练过程分为正向传播和反向传播两个过程。正向传播是将输入信息样本由输入层经过隐含层逐层处理并计算各单元的输出值。反向传播是指若由正向传播的输出值与期望输出值的误差没有达到一定

收稿日期: 2016-12-02

作者简介: 黄曦妮(1987-), 女, 大学本科, 助理工程师。

的标准,则逐层计算实际输出与期望值之间的差值,并根据此差值来调整权值的过程。各层之间传递函数一般为 S 型函数,该作用函数使得输出量为 0~1 之间的连续量,从而实现从输入到输出的任意非线性映射。对于第 p 个输入样本而言计算差值的误差函数见式(1)。

$$e_p = \frac{\sum_{n_3} (t_{in} - z_{in})^2}{2} \quad (1)$$

式中: e_p ——误差值;

i ——样本元素的个数;

t ——网络计算输出值;

z ——期望输出值。

BP 神经网络的计算流程图见图 2。

(1) 连接权值初始化

网络训练开始时一般用较小的随机数作为各层连接权值的初值。

(2) 各层输出值的计算

$$y_j = f_1 \left(\sum_{i=0}^{n_1} k_{1ji} x_i + b_i \right) \quad (j = 1, 2, \dots, 2n_1 + 1) \quad (2)$$

$$z_k = f_2 \left(\sum_{j=0}^{n_2} k_{2kj} y_j + b_j \right) \quad (k = 1, 2, \dots, n_3) \quad (3)$$

式中: x_i ——输入样本值;

y_j, z_k ——分别为隐含层的输出值和网络的实际输出值;

f_1, f_2 ——传递函数;

k_{1ji}, k_{2kj} ——分别为输入层与隐含层和隐含层与输出层之间的连接权值;

b_i, b_j ——分别为输入层和隐含层的阈值。

(3) 连接权值的修正

连接权值的修正采用负梯度法,连接各层连接权重修正量见式(4)、(5)。

$$\Delta k_{2kj} = -\eta \frac{\partial e}{\partial k_{2kj}} = \eta (t_k - z_k) f_2' y_j \quad (k = 1, 2, \dots, n_3; j = 1, 2, \dots, n_2) \quad (4)$$

$$\Delta k_{1ji} = -\eta \frac{\partial e}{\partial k_{1ji}} = \eta \sum_{k=1}^{n_3} (t_k - z_k) f_2' k_{2kj} f_1' x_i \quad (j = 1, 2, \dots, 2n_1 + 1) \quad (5)$$

式中: η ——学习速率。

(4) 阈值的修正

阈值也属于 BP 网络的参数,也需要进行修正,各层阈值修正量公式见式(6)、(7)。

$$\Delta b_j = \eta f_2' (t_n - z_n) \quad (6)$$

$$\Delta b_i = \eta \sum_{k=1}^{n_3} (t_k - z_k) f_2' b_j f_1' x_i \quad (7)$$

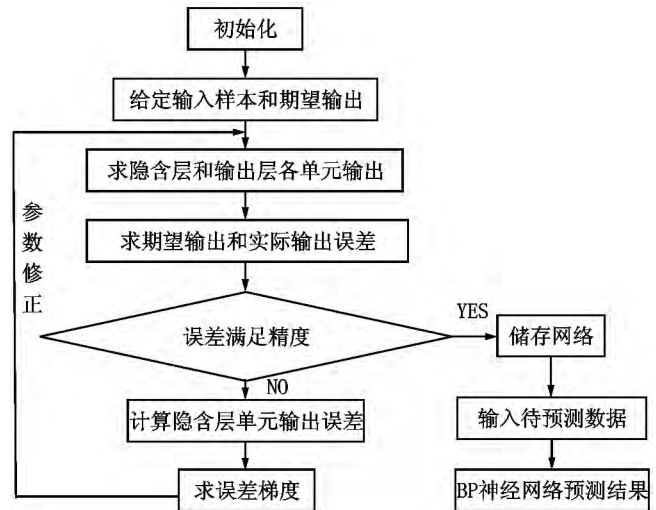


图 2 BP 神经网络计算步骤

2 BP 需水预测模型的建立与应用

下面以江西省廖坊水利枢纽工程灌区为例,采用 BP 神经网络方法对该灌区内的需水量进行预测,在预测基础上同其他方法计算预测的结果进行比较。

2.1 廖坊水利枢纽工程灌区概况

廖坊水利枢纽工程灌区(以下简称“廖坊灌区”)位于廖坊水利枢纽工程坝址下游、抚河流域的中下游地区,由东岸灌区、西岸灌区两部分组成。其东岸灌区位于东经 $116^{\circ}18' \sim 116^{\circ}42'$ 、北纬 $27^{\circ}41' \sim 28^{\circ}16'$ 之间;西岸灌区位于东经 $116^{\circ}20' \sim 116^{\circ}39'$ 、北纬 $27^{\circ}49' \sim 28^{\circ}00'$ 之间。廖坊灌区所灌农田主要分布于临川区、金溪县、东乡县及抚州高新区等四县(区)。灌区内多年平均降水量约为 $1\,730.0\text{ mm}$,降水量年内分配极不均匀,主要集中在 4~6 月份,占全年降水总量的 47.1%,7~9 月份是晚稻和其它农作物生长需水的关键时期,而此期间的降水量仅为年降水总量的 19.4%,约为 336.0 mm ,其同期的蒸发量则为 630.0 mm (D_{20})。随着经济的发展,灌区内工业用水和城市居民生活用水比重逐年增加,用水矛盾日趋突出,因缺乏灌溉水源,导致农业灌溉面积不断减少,而靠天赐水面积不断增加,形势十分严峻。

2.2 灌区需水量影响因子的确定

灌区需水量影响因子的确定是 BP 需水模型建立的重要组成部分,影响因子的确定决定了样本的特性,网络对由确定影响因子组成的训练数据以外样本的适应能力称为网络的泛化能力,泛化能力越高意味着网络的适应性越强。本文首先通过查阅抚州市年鉴以及相

关项目报告,初步确定了影响灌区需水量的几个指标,然后通过灌区现场调查以及向江西省水利规划设计研究院专家咨询,调研相关文献^[11]并结合灌区所处的经济、环境等特征进行分析,层层筛选,最终选定总人口数、工业产值、农田有效灌溉面积以及林牧渔业总产值这四个影响因子。灌区内总人口数、工业产值、农田有效灌溉面积、林牧渔业产值以及历年来的灌区需水量情况见表 1。各影响因子的用水量与总用水量的比值见表 2,它们对总用水量的贡献率达到了 90% 以上,可见选用这些影响因子作为 BP 网络的输入是具有—定代

表性的。

2.3 模型的建立与检验

本文应用灌区内 2003 ~ 2012 年的统计数据进行网络训练。BP 需水模型的建立与检验步骤如下:

(1) 输入数据的归一化,归一化处理用公式(8),归一化成果见表 3;

$$\frac{x_i - x_{\min}}{x_{\max} - x_{\min}} = x_{di} \tag{8}$$

式中: x_{di} 为归一化处理后的样本数据;

(2) 应用 MatLab 软件中的神经网络工具箱来建立

表 1 灌区影响因子以及灌区需水量

年份	总人口数/人			工业生产 总值/万元	林牧渔业生 产总值/万元	农田有效灌 溉面积/km ²	用水量 /亿 m ³
	城镇人口	乡村人口	合计				
2003	339 228	395 168	734 396	39 648	141 654	22. 11	3. 42
2004	341 635	397 973	739 608	56 315	159 450	22. 18	3. 77
2005	343 924	400 638	744 562	72 982	170 841	21. 50	3. 47
2006	346 151	403 233	749 383	114 998	184 574	20. 81	3. 49
2007	347 333	404 610	751 944	197 427	217 051	21. 34	4. 32
2008	348 528	406 002	754 530	258 136	272 562	21. 37	4. 53
2009	352 270	410 361	762 631	323 066	291 187	21. 50	5. 04
2010	352 869	411 058	763 927	467 252	320 765	21. 70	4. 84
2011	365 840	403 063	768 902	516 409	366 480	21. 49	5. 57
2012	379 357	391 711	771 068	652 206	398 044	22. 77	5. 05
2013	389 800	383 900	773 700	791 145	408 268	24. 05	5. 32

表 2 灌区影响因子占灌区总用水量的比值

年份	居民用水	工业生产	林牧渔业生产	农田灌溉
2003	0. 08	0. 05	0. 06	0. 74
2004	0. 08	0. 05	0. 06	0. 73
2005	0. 07	0. 05	0. 06	0. 74
2006	0. 07	0. 05	0. 06	0. 75
2007	0. 08	0. 06	0. 06	0. 73
2008	0. 07	0. 05	0. 06	0. 74
2009	0. 07	0. 05	0. 06	0. 75
2010	0. 07	0. 05	0. 05	0. 78
2011	0. 08	0. 05	0. 06	0. 78
2012	0. 07	0. 05	0. 06	0. 74
2013	0. 07	0. 05	0. 06	0. 75

网络,本研究中输入神经元有 4 个,网络隐含层的神经元有 9 个,网络中输出神经元有 1 个。首先用神经网络工具箱中的 newff() 函数编程来初始化网络,网络中间层的神经元传递函数采用正切型函数 tansig(),输出层神经元传递函数采用对数型函数 logsig()。模型参数根据一般经验确定为最大训练步数 h 为 3 000,学习速率 η 为 0. 1,动量因子 mc 为 0. 6,终止条件 goal 为 0. 001。根据设置的参数,用 trainlm 算法来训练输入的

样本数据,由于进行网络训练的初始权值是随机生成的,本研究运行多次采用训练结果最好的一次进行储存,便于之后的预测。经多次运行后取最优结果,可以得到训练误差曲线(plotperform)见图 3,图 3 表明网络在训练到第 6 次时就达到训练终止条件。

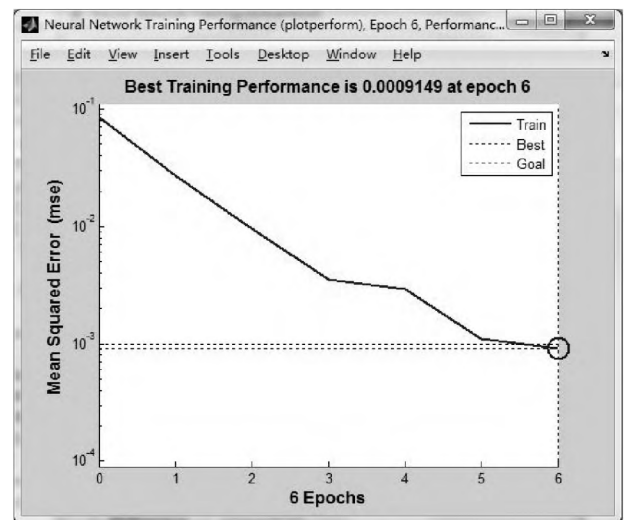


图 3 训练误差曲线

(3) 将 2013 年的统计数据作为检验模型的样本, 模拟出的结果为 0.881 3, 模拟结果反归一化后得到 2013 年的需水量为 5.31 亿 m^3 , 而 2013 年的实际需水量为 5.32 亿 m^3 , 可以看出与检验样本的期望值相比, 模型模拟值与它的相对误差为 1.9%, 模拟误差较小, 这说明建立 BP 需水模型可以用来预测该灌区的需水量。

2.3 廖坊灌区需水量预测

利用建立好的 BP 需水模型对廖坊灌区 2025 年的需水量进行预测, 各影响因子的值采用趋势法预测, 并与《廖坊水利枢纽工程灌区二期工程初步设计》报告用常规趋势法进行比较, 各预测方法比较结果见表 4。结果表明, 两种预测方法所得结果的相对误差为 1.6%, 预测结果比较合理。

表 3 归一化后的影响因子及需水量

年份	总人口数	工业生产 总值	林牧渔业 生产总值	农田有效 灌溉面积	用水量
2003	0.000	0.000	0.000	0.401	0.000
2004	0.133	0.022	0.067	0.424	0.161
2005	0.259	0.044	0.109	0.212	0.022
2006	0.381	0.100	0.161	0.000	0.035
2007	0.446	0.210	0.283	0.164	0.418
2008	0.512	0.291	0.491	0.170	0.515
2009	0.718	0.377	0.561	0.212	0.750
2010	0.751	0.569	0.672	0.274	0.658
2011	0.878	0.634	0.843	0.209	1.000
2012	0.933	0.815	0.962	0.604	0.757

(1) 总人口数的预测: 2013 年抚州市人口增长率为 6.93%, 采用指数函数增长方式, 见公式(9), 预测 2025 年的灌区内总人口数为 8.4×10^5 人。

$$P_n = P_0 e^{kn} \quad (n = 1, 2, \dots, n) \quad (9)$$

式中: P_n ——预测年份的人口数(人);

P_0 ——计算基准年的人口数(人);

e ——自然对数的底, 一般取 2.718;

k ——人口增长率;

n ——预测年数。

(2) 工业生产总产值的预测: 根据江西省统计年鉴, 近 3 年的抚州市工业生产总产值的年平均增长率为 16.33%。按式(10)预测 2025 年工业生产总产值为 4 859 111 万元。

$$Q_n = Q_{n-1} (1 + k) \quad (n = 2014, 2015, \dots, 2025) \quad (10)$$

式中: Q ——工业生产总产值(万元);

k ——工业生产总产值年平均增长率。

(3) 林牧渔业生产总产值的预测: 近 3 年的农林牧渔业生产总产值平均增长率为 8.5%, 采用与工业生产总产值相同的逐年增长法进行预测, 结果为 1 086 680 万元;

(4) 按照农业发展规划要求, 到 2025 年廖坊灌区所涉区域的粮食产量达 $88.30 \times 10^4 \text{t}$, 其主要措施是扩大种植面积, 充分利用目前暂未利用地, 提高单位面积产量, 对农业产业结构进行调整。故在此基础上农田有效灌溉面积将有所增加, 自 2011 年之后, 农田有效灌溉面积逐年增加量的平均值为 1.28 km^2 , 用公式(11)计算并结合实际情况即可得到 2025 年的农田有效灌溉面积为 33.49 km^2 。

$$S_n = S_{n-1} + k \quad (n = 2014, 2015, \dots, 2025) \quad (11)$$

式中: S_n ——预测年的有效灌溉面积(km^2);

S_{n-1} ——预测年的上一年的有效灌溉面积(km^2);

k ——有效灌溉面积逐年增加值(km^2)。

表 4 预测结果比较

预测方法	预测结果/亿 m^3	相对误差
BP 神经网络法	7.85	1.60%
常规趋势法	7.98	

3 结 论

本研究表明应用 BP 神经网络方法预测灌区需水量有很好的可信度和较强的预测能力, 此法在技术上是有效和可行的, 为制定灌区水资源配置规划方案提供了合理的方法, 开辟了新的途径。

参考文献:

- [1] 宰松梅, 郭冬冬, 温季. 基于最小二乘支持向量机的人民胜利渠灌区灌溉用水量预测[J]. 中国农村水利水电, 2009(12): 49~51.
- [2] 钱慕尧. 灰色预测在作物需水量推算中的应用[J]. 农田水利与小水电, 1993(10): 24~27.
- [3] 郭宗楼, 白宪台, 马学强. 作物需水量灰色预测模型[J]. 水电能源科学, 1995, 13(3): 186~191.
- [4] 张兵, 袁寿其, 成立, 等. 基于 L-M 优化算法的 BP 神经网络的作物需水量预测模型[J]. 农业工程学报, 2004, 20(6): 73~76.
- [5] 郭亚男, 吴泽宁, 高建菊. 基于主成分分析的支持向量机需水预测模型及其应用[J]. 中国农村水利水电, 2012(07): 76~82.
- [6] 臧冬伟, 陆宝宏, 朱从飞, 等. 基于灰色关联分析的 GA-BP 网络需水预测模型研究[J]. 水电能源科学, 2015(07): 39~43.
- [7] 崔东文. 基于相空间重构原理的遗传神经网络模型在城市需水预测

- 中的应用[J]. 水利水电科技进展, 2014, 34(1): 85~89.
- [8] 龙训建, 钱鞠, 梁川. 基于主成分分析的 BP 神经网络及其在需水预测中的应用[J]. 成都理工大学学报, 2010, 37(2): 216~220.
- [9] 焦斌, 叶明星. BP 神经网络隐层单元数确定方法[J]. 上海电机学院学报, 2013, 3(16): 114~117.
- [10] 赵振宁, 徐用慧. 模糊理论和神经网络的基础与应用[M]. 北京: 清华大学出版社, 1996.
- [11] 凌和良, 桂发亮, 楼明珠. BP 神经网络算法在需水预测与评价中的应用[J]. 数学的实践与认识, 2007(22): 42~47.

编辑: 唐少龙

Application of BP neural network method on water requirement forecasting in the irrigation area

HUANG Xini¹, ZHA Sihui²

(1. Jiangxi Provincial Water Conservancy Investment Group Corp., Nanchang 330006, China;
2. Hohai University, Nanjing 210098, China)

Abstract: Taking the irrigated area of Liaofang water conservancy project in Jiangxi province as an example, the main factors influencing the water demand were determined by referring to the yearbook, field investigation and expert consultation. The input samples of BP neural network were constructed according to the main influencing factors, and the water demand model of irrigated area was programmed by matlab software platform. The results show that BP neural network is an effective method to forecast the water demand of irrigation area in 2025 and BP neural network is successful in the application of water demand forecast in Liaofang irrigated area.

Key words: BP neural network; Water demand prediction; Irrigated area; Matlab;

翻译: 黄曦妮

江西省水利厅组织开展 2016 年度水利改革发展考核

根据《江西省委办公厅、省政府办公厅关于 2016 年市县科学发展综合考核评价实施意见》部署要求, 2017 年 1 月 9 日~13 日, 江西省水利厅派出由厅领导带队的 11 个省级考评组分赴全省各地开展水利改革发展省级考核。

考核主要包括水利建设、用水总量控制两方面。河长制落实及成效考核由省河长制办公室考核, 江西省水利厅负责河长制工作水利部分考核。水生态文明建设成效及粮食安全责任制考核与水利改革发展考核一并进行。

考评组通过听取汇报、查阅资料、现场检查等方式, 详细检查了各市县水利投资计划执行情况、党风廉政情况、水资源管理落实情况、河长制水利部分工作进展情况等。考核结束后, 考核组及时召开通气会, 反馈各地在水利改革发展推进、河长制落实、水生态文明建设等工作中取得的成效及存在的问题, 并提出针对性意见和建议。

市县科学发展综合考核评价体系是江西省委、省政府对市、县(市、区)进行考核考评的唯一体系。自 2011 年起, 我省将水利改革发展纳入对市、县政府考核评价体系。此项工作开展 6 年来, 江西省水利厅科学制定考核细则, 公平公正开展考核, 充分发挥考核工作的导向作用, 切实督促了各地加快水利改革发展工作进度, 全省水利系统形成了真抓实干、进位赶超的良好工作局面, 有力促进我省水利事业健康稳定发展。

(江西省水利厅办公室 李靖)