

灰色神经网络最优权组合模型预测城市需水量

蒋绍阶, 江崇国

(重庆大学 城市建设与环境工程学院, 重庆 400045)

摘要:需水量预测是一个大量数据指标和影响因素共同作用的复杂系统。目前以单一的模型预测为主,而这种预测方法仅能体现该系统的局部。针对这一情况,利用灰色模型和改进BP神经网络,建立最优权组合模型预测城市需水量,使用Matlab进行实例计算,并与其他预测方法比较。结果表明,该模型有较高的预测精度,优于单个模型,预测效果更优于其他方法。

关键词:需水量预测;灰色模型;改进BP神经网络;最优权组合模型

中图分类号: TU991.31 **文献标志码:** A **文章编号:** 1006-7329(2008)02-0113-03

Urban Water Demand Forecasting by Combining Improved BP Neural Network and Grey Model with Optimum Weight

J IANG Shao-jie, J IANG Chong-guo

(College of Urban Construction and Environmental Engineering, Chongqing University, Chongqing 400045, China)

Abstract: Prediction of water demand is a complex system affected by a mass of data and influencing factors together. Nowadays, most of the forecasting methods are single model ones. They reflect only part of the system. In view of this situation, a combined model is set up by use of grey model and improved BP neural network with optimum weight to forecast the urban water demand. It carries on the example computation by Matlab and compares the results with those by other prediction methods. It shows that the combined model has the high precision prediction. Its prediction results surpass those by the single model, even more surpass those by other ways.

Key words: prediction of water demand; grey model; improved BP neural network; combined model with optimum weight

城市需水量的量数值决定了给水工程系统所有组成部分的规模和布局,因此需水量预测是给水规划中最基本、最重要的工作之一。常用的预测方法为两类:时间序列预测法和解释性预测法^[1]。时间序列预测方法只依赖于历史观测数据来预测未来的需水量,其代表方法有时间序列分析法、灰色模型预测法等。解释性预测方法考虑到需水量变化受到独立的、具体的各个因素的影响,利用这些因素预测未来需水量,其代表方法有回归模型预测法、人工神经网络法等。

需水量变化是一个多种因素影响的复杂系统,单一的预测方法仅能体现系统的局部。如果建立不同方法的多个预测模型并通过适当的有效组合,则能充分利用各种信息提高精度预测,满足实际需要。

1 模型建立

1.1 预测模型选择

1.1.1 灰色模型

灰色系统理论将任何随机过程看作是在一定时空区域内变化的灰色过程,将随机变量看作是灰色量^[2];认为无规则的离散时空序列是在潜在的有规序列的某种表现,因而可以通过生成变换来弱化原始数列的随机性和离散情况,将无规序列变成有规序列。实践证明,灰色建模所需的信息较少,精度较高,能较好反映系统的实际状况^[3]。

GM(1,1)模型为单序列的一阶线性动态模型,是常用的一种灰色数列预测模型。该模型预测需水量的主要特点是所需自变量个数和历史数据个数都比较少。

* 收稿日期:2007-10-21

作者简介:蒋绍阶(1956-)男,湖南祁阳人,副教授,博士,主要从事给水处理研究,(E-mail) szhjzx@126.com。

1.1.2 改进的BP神经网络

神经网络是由大量简单的神经元相互连接构成的复杂网络。神经元是对生物神经元的简化和模拟。神经网络具有模式识别能力、自组织、自学习、非线性等独特的优点^[4]。已经证明三层的前向BP神经网络可以任意的精度逼近任意非线性函数^[5],而且运用神经网络不需要模型,只要有输入和目标输出就可以了,所以用神经网络来做城市需水量的预测是十分有效的。

改进的BP神经网络,隐含层的激活函数仍采用S型函数中的Sigmoid函数: $f(x) = \frac{1}{1+e^{-x}}$ 。通过附加动量法修正网络的权值,使网络避免陷入局部极小值。本文采用动量梯度下降方法对权值和阈值进行调整,调整值 dW 由动量因子 mc 、前一次学习时的调整值 dW_{prev} 、学习速率 lr 和梯度 gW 共同确定^[6]。

$$dW = mc \cdot dW_{\text{prev}} + (1 - mc) \cdot lr \cdot gW$$

1.2 组合模型建立

最优权组合模型的基本思路是:先优选几个统计模型,然后选取一组最优的权重将它们线性组合起来,从而得到组合模型。由此可见,建立组合模型的关键就是合理地确定权系数向量。

设优选了 m 个统计模型 $\hat{y}_i (i = 1, 2, \dots, m)$, 由它们组成组合模型 $Y = \varphi(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m)$, 组合模型中各单个模型的权重组成向量 $\omega^T = [\omega_1, \omega_2, \dots, \omega_m]$, 并取 $\sum_{i=1}^m \omega_i = 1$, 从而构成下列组合模型^[7]:

$$Y = \sum_{i=1}^m \omega_i \hat{y}_i = \omega_1 \hat{y}_1 + \omega_2 \hat{y}_2 + \dots + \omega_m \hat{y}_m \quad (1)$$

记 $y^T = (y_i | i = 1, 2, \dots, n)$ 为实测值向量;

$\hat{y}_i = (\hat{y}_{it} | t = 1, 2, \dots, n)$ 为第 i 个模型的拟合值

向量, $i = 1, 2, \dots, m$;

$Y = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m]$ 为各单模型的拟合值矩阵;

$e_i = (e_{it} = y_i - \hat{y}_{it} | t = 1, 2, \dots, n)$ 为第 i 个模型的拟合残差向量;

$E = [e_1, e_2, \dots, e_m]$ 为单模型的拟合残差矩阵;

组合模型的拟合值向量和拟合残差向量可表示为:

$$\hat{y}^T = \omega^T Y^T \quad (2)$$

$$e^T = \omega^T E^T \quad (3)$$

依据残差平方和最小的准则,可构造以下数学规划模型:

$$\begin{aligned} \min Q &= e^T e = [\omega^T E^T] \cdot [\omega^T E^T]^T \\ &= \omega^T (E^T E) \omega = \omega^T B \omega \end{aligned} \quad (4)$$

$$\text{s.t. } R^T \omega = 1 \quad (5)$$

式中 $R^T = (1, 1, \dots, 1)$ 为元素全为1的 rn 维列向量;
 $E = E^T E$ 。

对式(4), (5)运用拉格朗日乘法求解,可求得模型唯一解:

$$\begin{aligned} \omega &= \frac{E^{-1} R}{R^T E^{-1} R}, \\ \min Q &= \frac{1}{R^T E^{-1} R} \end{aligned} \quad (6)$$

利用式(6)即可求出组合模型中各单个模型的最优权重比与最小残差平方和。这时按式(1)构成的组合模型将优于任一个单个模型的拟合精度。

2 应用实例

为了验证上述方法和模型的实用程度,对收集的南方某城市需水量预测基础数据资料(如表1)进行分析与计算。

表1 1996~2005年人口、生产总值环比发展速度、用水量统计资料表

年份	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
年末常住人口/万人	358.48	379.64	394.96	405.14	432.94	468.76	504.25	557.41	597.55	(623.51)
生产总值环比速度/%	116.4	116	114.5	114.1	114.2	113.2	115	119.2	117.3	115
用水量/万 m ³	64 743	70 194	79 596	86 560	92 068	97 334	108 070	122 795	129 438	139 487

注:由于2005年某市统计人口数还未公布,所以使用曲线拟合法预测2005年人口:623.51万人(拟合方程: $y = -0.1535x^4 + 3.1626x^3 - 19.0519x^2 + 58.5354x + 316.0761$)。

2.1 模型计算

灰色预测法。取1996~2005年用水量数据,建立GM(1,1)模型,求得发展灰数 $a = -0.085701$, 内生控制灰数 $u = 62778$, 进一步计算得预测公式^[8]:

$$\hat{x}^{(0)}(k) = 65480.1549e^{0.0857(k-1)}, (k = 2, \dots, m)$$

还原可得拟合预测值(见表2)。

改进BP神经网络预测法。取某市水量变化的主要影响因子,1996~2005年人口和GDP增长率作为基础数据,将其极差规格化后作为输入向量,将规格化后的历年用水量数据作为输出向量,指定精度 $\varepsilon = 0.0001$ 。通过Matlab试算^[6]后,建立2个输入层神经元、3个隐藏层神经元和1个输出层神经元的3层神经网络,得预测值(见表2)。

表2 需水量拟合数据与原始数据误差比较表

年份	实际 用水量 /万 m ³	灰色模型		BP神经网络		组合模型	
		预测值 /万 m ³	误差 /%	预测值 /万 m ³	误差 /%	预测值 /万 m ³	误差 /%
1996	64 743	65 480	1.14	65 338	0.92	65 419	1.04
1997	70 194	71 339	1.63	72 153	2.79	71 689	2.13
1998	79 596	77 723	-2.35	78 149	-1.82	77 906	-2.12
1999	86 560	84 677	-2.18	83 435	-3.61	84 144	-2.79
2000	92 068	92 254	0.20	91 983	-0.09	92 138	0.08
2001	97 334	100 509	3.26	101 024	3.79	100 730	3.49
2002	108 070	109 502	1.33	109 396	1.23	109 457	1.28
2003	122 795	119 301	-2.85	120 326	-2.01	119 741	-2.49
2004	129 438	129 976	0.42	130 378	0.73	130 149	0.55
2005	139 487	141 606	1.52	140 191	0.50	140 998	1.08

上述两组预测值组成拟合值向量,与对应实际值运算,使用 Matlab 软件求得拟合残差矩阵 E ,按式(6)计算得权重向量 $\omega^T = [0.570\ 4\ 0.429\ 6]$ 。其中灰色预测的权重系数为 0.570 4,神经网络预测的权重系数为 0.429 6,进一步计算出预测值(见表 2)。

2.2 模型精度检验

由表 2 可知,各方法所预测数据与实际值的相对误差绝对值均小于 4%,均方差 σ 分别是 1951、1972、1923,所以,组合预测的拟合精度、预报效果要优于灰色预测和神经网络预测。

2.3 组合模型与用水性质预测值比较

用水性质预测法是在土地及水资源承载力的限定条件下,通过对目前大量资料及现状进行调研和分析,

辅以国内外常用的各种指标校核预测值的方式,给出现阶段条件下合理的预测结果。此法具有多元预测的特点,是现有方法中较全面、系统的方法。根据某市特点,分析从统计局和水务局公报中搜集的数据指标(见表 3),按居民生活、工业、建筑、三产等四类主要用户分别对 2002~2004 年需水量进行预测(见表 4)。

表 3 2002~2004 年社会经济发展指标表

年份	常住人口 /万人	GDP /亿元	GDP 比重/ %			
			一产	二产		三产
				工业	其他	
2002	504.25	2 240	0.8	47.9	7.3	44
2003	557.41	2 861	0.6	52.4	6.5	40.5
2004	597.55	3 423	0.3	48.1	3.5	48.1

表 4 2002~2004 年用水性质法需水量预测表

年份	居民生活		工业		建筑业		三产		城市需水量
	定额	需水量	万元	GDP 耗水量	需水量	人均定额	需水量	增加值定额	
	L/ P. d	亿 m ³	m ³ / 万元	亿 m ³	m ³ / 人	亿 m ³	m ³ / 万元	亿 m ³	
2002	240	4.4	59.5	6.4	10	0.7	17	1.7	13.1
2003	275	5.6	51.5	7.7	10	0.7	16	1.9	15.9
2004	280	6.1	46.8	7.7	10	0.8	15	2.5	17.1

由表 4 可知,2002~2004 年预测值与实际值的相对误差分别为 21.5%、29.4%、31.8%,其预测结果比组合模型预测的结果误差大,可见采用最优权组合模型可以提高需水量预测精度。

2.4 预测结果

根据“十一五”总体规划,2010 年全市生产总值达到 9 000 亿元,五年平均增长 13%左右,另外人口采用水利部 2005 年 5 月预测的结果:2010 年常住人口 750 万人。灰色预测预测某市近期需水量为 21.7 亿 m³,神经网络预测为 25.9 亿 m³,所以,最优权组合模型预测 2010 年某市需水量为 23.5 亿 m³。

3 结 论

灰色模型和 BP 神经网络都是较好的需水量预测方法,在预测中都有较高的准确度。通过最优权模型组合,不仅克服自身缺点,而且使预测结果更可靠、准确,有利于后期决策。

参考文献:

[1] 白雪华,郭旭颖.改进的 RBF 网络在区域需水量预测中的应用[J]. 青岛建筑工程学院学报,2005,26(3):87-89.
BAI Xue-hua, GUO Xu-ying. Applications of water demand prediction based on improved RBF neural network [J]. Journal of Qingdao Institute of Architecture and Engineering, 2005, 26(3): 87-89.

[2] 蒋承仪.灰色马尔柯夫预测模型[J]. 重庆建筑大学学报,1996,18(3):116-122.
JIANG Cheng-yi. Grey-Markov forecasting model [J]. Journal of Chongqing Jianzhu University, 1996, 18(3): 116-122.

[3] 潘敏,范庆来.灰色模型在城市用水量预测中的应用[J]. 浙江水利科技,2004,(1):21-22.
PAN Ming, FAN Qing-lai. Applications of water demand prediction based on grey model [J]. Zhejiang Science and Technology of Water Conservancy, 2004, (1): 21-22.

[4] 胡文发.基于 BP 算法的国际工程项目政治风险评价模型[J]. 重庆建筑大学学报,2006,28(4):98-100.
HU Wen-fa. B Palgorithm evaluation model of political risk in international construction projects [J]. Journal of Chongqing Jianzhu University, 2006, 28(4): 98-100.

[5] 张雪飞,郭秀锐,等.BP 神经网络法预测唐山需水量[J]. 安全与环境学报,2005,5(5):95-98.
ZHANG Xue-fei, GUO Xiurui. Prediction of urban water demand in Tangshan city with B Pneutral network method [J]. Journal of Safety and Environment, 2005, 5(5): 95-98.

[6] 许东,吴铮.基于 MATLAB 6.x 的系统分析与设计——神经网络[M]. 西安:西安电子科技大学出版社,2002.

[7] 田斌,任德记,何薪基.非线性组合优化模型在水工建筑物位移监控中的应用[J]. 水电自动化与大坝监测,2004,28(4):55-58.
TIAN Bin, REN De-ji, HE Xin-ji. Application of nonlinear composite optimized model to displacement monitoring of hydro-structure [J]. Hydropower Automation and Dam Monitoring, 2004, 28(4): 55-58.

[8] 陈勇. GM(1,1) 灰色模型的程序实现[J]. 固原师专学报(自然科学),2005,26(3):31-33.
CHEN Yong. Program implementation of GM(1,1) grey model [J]. Journal of Guyuan Teachers College (Natural Sciences), 2005, 26(3): 31-33.

(编辑 王秀玲)