

doi:10.11835/j.issn.1674-4764.2017.05.019



混沌局域法与神经网络组合供水量预测

孙晓婷, 刘年东, 杜坤, 周明, 任刚红

(昆明理工大学 建筑工程学院, 昆明 650500)

摘要:城市供水量是非线性、非平稳时间序列,组合预测模型能获得更高精度预测结果。通过深入分析混沌局域法与神经网络预测模型特点,提出了一种新的组合预测模型。首先,应用混沌局域法对城市日供水量进行初预测,然后,应用神经网络对预测结果进行修正。由于所提出的组合模型利用了混沌局域法及神经网络进行优势互补,能同时提高预测精度与计算效率。为验证所提出组合预测模型的可行性,采用某市 7 a 实测供水量数据,对混沌局域法、BPNN、RBF 及 GRNN 神经网络 4 种单一预测模型及相应的 3 种组合模型预测精度进行定量分析,结果表明,组合预测模型精度都高于对应单一预测模型,混沌局域法与 GRNN 神经网络组合模型预测精度最高,且运算时间远低于单一神经网络模型运算时间。

关键词:混沌局域法;神经网络;组合模型;日供水量预测

中图分类号:TP183 **文献标志码:**A **文章编号:**1674-4764(2017)05-0135-05

Water supply forecasting based on the combination of chaotic local-region method and neural network

Sun Xiaoting, Liu Niandong, Du Kun, Zhou Ming, Ren Ganghong

(Faculty of Civil Engineering and Mechanics, Kunming University of Science and Technology, Kunming 650500, P. R. China)

Abstract: Urban water supply is a nonlinear and non-stationary time series, and the combination forecasting model can get more accurate results. Through in-depth analysis of chaotic local-region method and neural network prediction model, this paper puts forward a new combination forecasting model, which uses chaotic local-region method to make a preliminary forecast for urban daily water supply, and then the prediction result is updated by neural network. The proposed combined model makes use of complementary advantages of the chaotic local-region method and the neural network, improving synchronously the accuracy and computational efficiency of the prediction results. To verify the proposed model, the prediction accuracy of the four single prediction models of Chaotic local-region method, BPNN, RBF and GRNN neural network and three corresponding combined models are analyzed quantitatively using seven years water supply data. The results show that combination forecasting model is of higher accuracy than single prediction model, and chaotic local-region method plus GRNN neural network combination model has

收稿日期:2017-02-05

基金项目:国家自然科学基金(51608242)

作者简介:孙晓婷(1994-),女,主要从事市政工程研究,(E-mail)2425285836@qq.com。

杜坤(通信作者),男,博士,(E-mail) dukun_cq@foxmail.com。

Received:2017-02-05

Foundation item:National Natural Science Foundation of China(No. 51608242)

Author brief:Sun Xiaoting(1994-),main research interest:municipal engineering,(E-mail)2425285836@qq.com.

Du Kun(corresponding author),PhD,(E-mail) dukun_cq@foxmail.com.

highest accuracy with much lower computation time than single neural network predication model.

Keywords: chaotic local-region method; neural network; combination model; daily water supply forecast

城市供水量预测能辅助供水调度,提高水厂管理水平与生产效率,一直是学者们关注的重点课题^[1-2]。供水量预测模型可分为传统预测模型和基于新技术预测模型^[3],传统模型需对数据序列性质进行假设,例如,平稳性假设或周期性假设,若假设不合理,得出的预测模型则会严重失真;基于新技术的预测模型通过非线性、自适应学习方法构建模型,能克服传统预测模型缺点。如Tiwari等^[4]提出了一种基于小波技术的神经网络供水量短期预测模型,结果表明,其预测精度比传统 ARIMA、ARIMAX 和 WNN 方法高。Bai等^[5]分析了供水量序列的混沌特性,利用自适应混沌粒子群优化 RVM 模型参数,提出一种多尺度的 RVM 供水量预测组合模型。陈敏等^[6]根据混沌理论计算重构相空间嵌入维数,用于确定 BP 神经网络隐藏层节点个数,提高了预测精度。

从信息利用角度来看,单一预测模型只能利用数据部分有效信息,仅能从一个侧面刻画数据序列规律,具有一定局限性;组合预测模型通过优势互补,能更大程度挖掘数据信息,可望获得更高精度预测结果。神经网络与混沌理论模型作为目前最广泛使用的两种新技术预测模型,都展现出较高的预测精度,但针对二者的组合预测模型研究较少。通过深入分析,笔者发现,基于混沌理论的预测模型^[7],例如:全域法、局域法等,其特点是在海量数据样本中迅速挖掘时间序列总体趋势,但对局部细节预测能力较差。相较而言,神经网络模型^[8-9]非线性拟合能力更强,能更准确追踪局部细节,但当样本数据量较大时,神经网络预测模型存在训练时间长、收敛慢且预测结果不确定的缺点。鉴于此,提出了一种新的组合预测方法,首先,利用混沌预测模型对数据序列总体趋势进行初预测,然后,应用小样本训练神经网络对预测结果进行局部修正,实现二者的优势互补,进而能同时提高预测结果精度与计算效率。考虑混沌局域模型能较好处理实际中存在噪声的预测问题,采用加权一阶局域法进行初预测,重点研究了其与不同类型神经网络的组合预测效果,并采用某市的实测供水量数据对单一模型及组合模型的预测精度、运算时间进行定量分析。

1 基于混沌理论的加权一阶局域预测法

混沌局域预测法将相空间运行轨迹的最后一点作为基准点,通过寻找历史相点中最近的若干相点作为运动趋势参考点,预测基准点走向,其步骤可概括为:1) 重构相空间,包括嵌入维与时间延迟的确定;2) 邻近相点选取;3) 选择拟合方法进行预测计算。重构相空间^[10]是将一维混沌时间序列映射到高维的空间,目的是恢复有规律的吸引子,从而使蕴藏在时间序列中的信息显露。设混沌时间序列 $\{x_1, x_2, \dots, x_N\}$, 相空间重构为

$$X(t_i) = [x(t_i), x(t_i + 2\tau), \dots, x(t_i + (m-1)\tau)] \\ (i = 1, 2, \dots, N - (m-1)\tau) \quad (1)$$

式中: τ 为时间延迟; m 为嵌入维数, τ 与 m 的取值决定了相空间状态。常用确定 τ 与 m 的方法有:自相关法、互信息法、G-P 法、Cao 法以及 C-C 法等。

在重构相空间后,需要选取邻近点,即确定基准相点的运动趋势参考点。研究表明^[11],邻近相点对预测精度有直接影响,邻近相点个数选择过少,则历史运动趋势可能无法被充分反映,但选取过多的邻近相点,不仅会加大计算量,还可能引入与运动趋势不相关的相点,降低预测精度。本文以每个相点到基准点的欧式距离作为评判指标,采用 BIC 信息准则^[12]确定邻近点个数。设由邻近相点构成的矩阵为 \mathbf{X}_M , 邻近相点的下一步演化点矩阵为 \mathbf{X}_{M+1} , 则基于混沌理论的一阶局域预测模型为

$$\mathbf{X}_{M+1} = a\mathbf{e} + b\mathbf{X}_M \quad (2)$$

式中: $\mathbf{e} = [1 \ 1 \ \dots \ 1]_m^T$; a 和 b 为拟合参数。若要计算基准相点的第 n 步预测值,采用邻近相点的第 n 步演化点作为参考点即可,即将上述模型中的 \mathbf{X}_{M+1} 替换为 \mathbf{X}_{M+n} 。

采用加权一阶局域预测模型时,其权系数可根据邻近点到基准点的欧式距离确定,具体如下:设基准点 \mathbf{X}_M 的邻近相点集为 $\{\mathbf{X}_{M_i}, i = 1, 2, \dots, q\}$, $\{\mathbf{X}_{M_i}\}$ 中每个相点到基准点的欧式距离为 d_i , 令 d_i 中的最小值为 d_{\min} , 定义 $\{\mathbf{X}_{M_i}\}$ 的权值为 P_i , 则

$$P_i = \frac{e^{-a(d_i - d_{\min})}}{\sum e^{-a(d_i - d_{\min})}} \quad (3)$$

式中: α 为权重调节系数,一般 $\alpha=1$ 。根据加权最小二乘法, a 、 b 能计算为

$$[a \quad b]^T = (A^T P A) A^T P Y \quad (4)$$

式中: $Y = X_{M+1}$; $A = [e \quad X_M]$; 权重矩阵 P 为对角矩阵, 对角元素为 P_i 。在求解出参数 a 、 b 后, 将基准点带入式(4)即可得预测值。

2 神经网络及其与加权一阶局域法的组合预测模型

人工神经网络是一种能对信息进行分布式并行处理的数学模型, 其最大特点是具有自学习和自适应能力^[13]。根据已有文献, BP、RBF 及 GRNN 神经网络发展相对成熟, 被广泛用于解决各类预测问题。BP 神经网络^[14-15], 即反向传播神经网络, 采用输出与输入之差逆向反馈校正的方法使实际输出不断逼近期望输出。它也是是目前理论最为完备的神经网络, 但它要事先确定网络的结构, 参数调整复杂, 人为干扰因素较强, 训练好的网络, 在给它新的时序时, 就需要重新训练。RBF 网络^[16], 即径向基神经网络, 利用用 RBF 作为隐单元的“基”构成隐藏空间, 将低维的模型输入数据变换到高维空间内, 使得在低维空间的线性不可分问题在高维空间线性可分。广义回归神经网络^[17-18] (GRNN) 也是一种径向基神经网络, 仅需要调节一个参数, 因此, 人为干扰因素较小。在小数据量的情况下, 预测效果也很好且运算时间更短。

将加权一阶局域法与上述 3 种神经网络分别进行组合, 利用某市实测供水量数据评判单一模型及组合模型的预测精度, 具体步骤为:

1) 将时间序列样本 X 分为两部分, 即前 $N - m$ 个样本及后 m 个样本, m 为相空间嵌入维数, 预测步距为 i ;

2) 将前 $N - m$ 个时间序列作为样本, 应用加权一阶局域法预测后 m 个样本, 并计算预测误差即 $e_m = X_m - X'_m$ 。同时再计算出 N 以后的 i 步预测值 X'_{N+i} ;

3) 应用 2) 中所得的 m 个预测误差样本训练神经网络, 对加权一阶局域法预测误差 e_m 进行预测, 得到预测结果为 e'_{N+i} ;

4) 应用加权一阶局域法预测后 $N + m$ 个样本, 并根据 3) 所得的预测误差修正预测结果为 $X_{N+i} = e'_{N+i} + X'_{N+i}$, 其流程如图 1 所示。

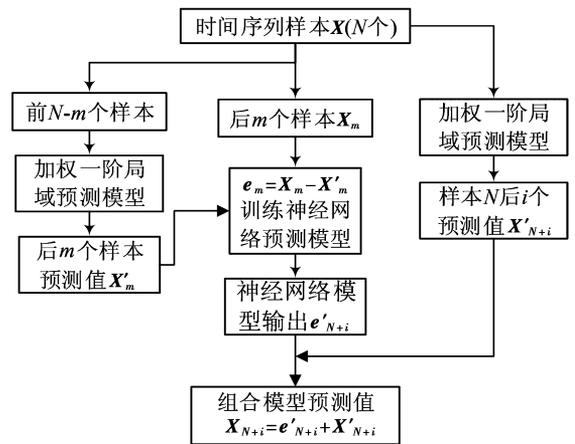


图 1 加权一阶局域法与神经网络的组合预测流程图
Fig. 1 Combined forecasting flow chart of weighted first order local area method and neural network

3 案例分析

选取某市水厂 2000 年 1 月至 2006 年 12 月的日供水验证预测模型精度。为消除年供水量时间序列的季节性和趋势性、减少噪声影响^[19], 选取 2000~2006 年 1 月的 217 个日供水量作为时间序列样本, 其中 210 个日供水量数据作为参考样本, 7 个日供水量数据作为验证样本, 对于其它月份的预测可按此方法进行处理。采用互信息法计算得 $\tau=7$, 根据文献[20]推荐的多嵌入维法计算得 $m=10$, 再依据 BIC 信息准则选取邻近相点 $K=7$ 。首先采用加权一阶局域预测法、BP、RBF 及 GRNN 神经网络预测模型对日供水量进行单独预测, 然而将 3 种神经网络分别与加权一阶局域法进行组合预测, 总体预测趋势见图 2, 局部细节见图 3。

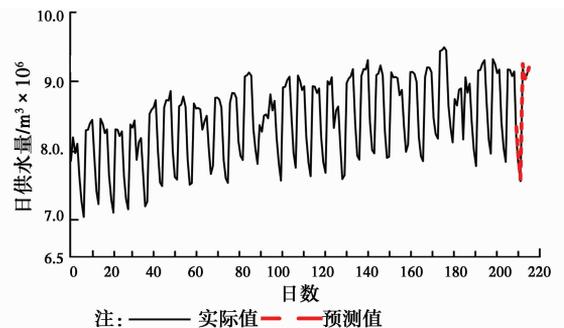


图 2 日供水量总体预测趋势

Fig. 2 General forecast trend of daily water supply

如图 2、图 3 所示, 7 种方法都能较好预测日供水量总体趋势, 而对局部细节预测有一定差异。对预测结果按精度排序, 将平均绝对误差 (MAPE) 最小的置于首位, 如表 1 所示。混沌局域法与 GRNN

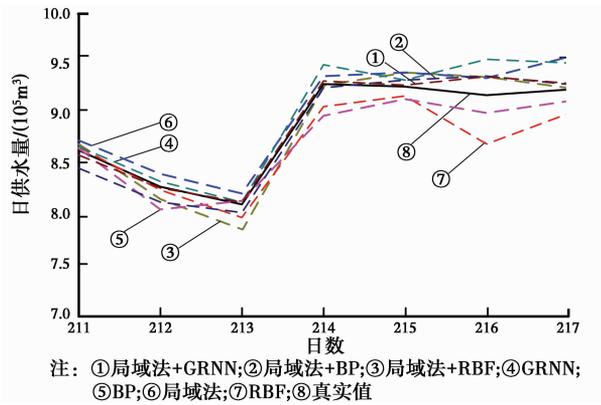


图 3 日供水量局部预测细节

Fig. 3 Local forecast details of daily water supply

组合模型预测精度最高,单一的 RBF 神经网络预测精度最低。此外,组合预测模型的精度都高于相应的单一模型,由于采用小样本训练神经网络,其运算时间仅为 3 s,达到了同时提高预测结果精度及计算效率的目的。

表 1 供水量预测误差及运算时间

Table 1 Forecasting error and operation time of daily water supply

预测模型	供水量预测误差/%				运算时间/s
	211	212	213	214	
局域法+GRNN	-0.71	-0.21	0.35	0.42	3
局域法+BP	-2.67	-2.44	-1.32	-0.50	4
局域法+RBF	0.95	-1.98	-4.10	-0.15	3
GRNN	0.58	0.84	0.35	2.64	90
BP	0.30	-3.60	0.58	-4.31	90
混沌局域法	1.50	2.03	1.78	1.08	78
RBF	0.49	-0.53	-2.12	-3.08	3

预测模型	供水量预测误差/%			MAPE
	215	216	217	
局域法+GRNN	0.21	2.57	0.85	0.76
局域法+BP	0.90	2.64	0.71	1.60
局域法+RBF	1.93	2.51	0.18	1.68
GRNN	0.90	4.95	3.64	1.99
BP	-1.70	-2.45	-1.60	2.08
混沌局域法	1.94	2.36	4.46	2.16
RBF	-1.28	-6.74	-3.30	2.51

4 结论

供水量预测对提高水厂管理水平具有重要意义,本文提出了一种新的组合预测方法以提高供水量预测精度,其利用混沌加权一阶局域法对供水量

进行初预测,然后利用神经网络对预测结果进行修正。采用某市 7 年实测的供水量数据对不同模型预测结果进行评判,得到了如下结论:

1)对单一模型,神经网络预测精度普遍高于混沌加权一阶局域法,其中 GRNN 神经网络预测精度最高,BP 神经网络预测精度次之,但神经网络预测模型存在训练时间长、预测结果不确定的缺点。

2)组合预测模型精度都高于对应单一预测模型,其中加权一阶局域法与 GRNN 神经网络的组合模型预测精度最高,加权一阶局域法与 RBF 神经网络的组合模型预测精度最低,但仍高于所有单一预测模型。由于采用了小样本训练神经网络模型,其运算时间短且精度更高。

参考文献:

[1] YASAR A, BILGILI M, SIMSEK E. Water demand forecasting based on stepwise multiple nonlinear regression analysis [J]. Arabian Journal for Science and Engineering, 2012, 37(8): 2333-2341.

[2] 孙月峰,闫雅飞,张表志. 基于 T-S 模型的模糊神经网络城市需水量预测方法研究[J]. 安全与环境学报, 2013, 13(2): 136-139.

SUN Y F, YAN Y F, ZHANG B Z. Fuzzy neural network method based on the T-S model for the urban water demand forecasting and its application [J]. Journal of Safety and Environment, 2013, 13(2): 136-139. (in Chinese)

[3] PAI Y, PU W, CHUAN L, et al. Dynamic forecast of daily urban water consumption using variable-structure support vector regression model[J]. Journal of Water Resource Planning and Management, 2014, 140: 1943-5452.

[4] TIWARI M K, ADAMOWSKI J. Urban water demand forecasting and uncertainty assessment using ensemble wavelet-bootstrap-neural network models [J]. Water Resources Research, 2013, 49(10): 6486-6507.

[5] BAI Y, WANG P, LI C, et al. A multi-scale relevance vector regression approach for daily urban water demand forecasting [J]. Journal of Hydrology, 2014, 517: 236-245.

[6] 陈敏,邓英伟. 基于混沌理论的城市用水量预测研究[J]. 计算技术与自动化, 2009, 28(1): 88-90.

CHEN M, DEN Y W. Forecast research of urban water consumption based on chaos theory [J]. Computing Technology and Automation, 2009, 28(1): 88-90. (in Chinese)

- [7] 韩敏. 混沌时间序列预测理论与方法[M]. 北京:中国水利水电出版社,2007:28-39.
HAN M. Prediction theory and method of chaotic time series [M]. Beijing: China Water & Power Press, 2007:28-39. (in Chinese)
- [8] 史峰,王小川,郁磊. MATLAB神经网络30个案例分析[M]. 北京:北京航空航天大学出版社,2011:73-80.
SHI F, WANG X C, YU L. 30 case analysis of MATLAB neural network [M]. Beijing: Beihang University Press,2011:73-80. (in Chinese)
- [9] 林海军,滕召胜,迟海,等. 基于多RBF神经网络的汽车衡误差补偿[J]. 湖南大学学报(自然科学版),2010,37(5):40-44.
LIN H J, TENG Z S, CHI H, et al. Error compensation for truck scale based on multiple RBF neural networks [J]. Journal of Hunan University (Natural Sciences), 2010,37(5):40-44. (in Chinese)
- [10] TAKENS F. Determining strange attractors in turbulence [J]. Lecture Notes in Mathematics, 1981, 898: 361-381.
- [11] 刘年东,杜坤,周明,等. 局域法邻近点选取对降雨量预测精度影响研究[J]. 给水排水, 2016(Sup1):285-288.
LIU N D, DU K, ZHOU M, et al. Study on the influence of select the local-region method nearest neighbours on rainfall forecasting accuracy[J]. Water & Wastewater and Drainage, 2016(Sup1):285-288. (in Chinese)
- [12] CARLOS C, MUÑOZ-COLLAZOS H, URBANO-MUÑOZ R. Clustering of web search results based on the cuckoo search algorithm and Balanced Bayesian Information Criterion [J]. Information Sciences, 2014, 281:248-264.
- [13] TAORMINA R, CHAU K, SETHI R. Artificial neural network simulation of hourly 4 groundwater levels in a coastal aquifer system of the 5 Venice lagoon [J]. Engineering Applications of Artificial Intelligence, 2012, 25 (8): 1670-1676.
- [14] KARSOLIYA S. Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture [J]. International Journal of Engineering Trends and Technology, 2012, 3(6): 714-717.
- [15] 戴张俊,余飞,陈善雄,等. 淮北平原地基分层与承载力的BP网络预测方法[J]. 土木建筑与环境工程, 2013, 35(3):18-24.
DAI Z J, YU F, CHEN S X, et al. Stratification and bearing capacity prediction method based on BP neural network for foundation in Huaibei Plain [J]. Journal of Civil, Architectural & Environmental Engineering, 2013,35(3):18-24. (in Chinese)
- [16] TSEKOURASA G E, TSIMIKASB J. On training RBF neural networks using input-output fuzzy clustering and particle swarm optimization [J]. Science Direct Fuzzy Sets and Systems,2013, 221:65-89.
- [17] 戴虹,钱晋武,张震,等. GRNN在肌电预测踝关节运动中的应用[J]. 仪器仪表学报, 2013,34(4):846-851.
DAI H, QIAN J W, ZHANG Z, et al. Application of GRNN in ankle movement prediction based on surface electromyography [J]. Chinese Journal of Scientific Instrument, 2013,34 (4):846-851. (in Chinese)
- [18] FIRATA M, TURANB M E, YURDUSEV M A. Comparative analysis of neural network techniques for predicting water consumption time series [J]. Journal of Hydrology, 2010, 384 : 46-51.
- [19] 赵莹,卢文喜,罗建男,等. 基于改进时间序列分析法的镇赉地区地下水位动态分析[J]. 水利学报,2013,44(11):1372-1379.
ZHAO Y, LU W X, LUO J N, et al. Dynamic analysis of groundwater table in Zhenlai county of Jilin province based on the improved time series analysis method [J]. Shuili Xuebao, 2013,44(11):1372-1379. (in Chinese)
- [20] 刘道文,忽海娜. 基于多嵌入维加权一阶局域法煤矿灾害混沌时间序列预测[J]. 煤炭技术,2012,31(8):89-91.
LIU D W, HU H N. Chaotic time series prediction of coal mine disasters based on multi-dimensional weighted first-order embedded local-region method [J]. Coal Technology, 2012,31(8):89-91. (in Chinese)

(编辑 胡玲)