

DOI: 10. 20040/j. cnki. 1000-7709. 2023. 20230267

基于神经网络的南京市用水量预测

王浏琳^{1a,1b}, 陈家栋², 张方敏^{1a,1b}

(1. 南京信息工程大学 a. 气象灾害预报预警与评估协同创新中心; b. 应用气象学院江苏省农业气象重点实验室, 江苏 南京 210044; 2. 江苏省水文水资源勘测局南京分局, 江苏 南京 210008)

摘要: 分析南京市用水量现状,合理预测南京市用水量,掌握未来用水需求,对南京市水资源的长远规划和配置具有重大意义。对南京市 2009~2019 年用水量分析表明,工业、农业用水在南京市总用水量中比重较大,对总用水量的变化起着至关重要的作用。其次,采用灰色预测 GM(1,1)模型与 Elman 神经网络的组合模型预测南京市各区及全市总用水量。结果表明,灰色 Elman 神经网络模型对 2009~2019 年南京市全市总用水量的预测效果良好,预测的相对误差均小于 3.5%,预测结果的多年平均相对误差为 1.55%;在南京市 2019 年各区用水量预测中,预测结果的相对误差均小于 8.5%,效果较好。可见所用模型能准确地预测南京市用水量,对有效控制区域用水量,实现“四水四定”原则具有重要意义。

关键词: 机器学习;灰色 Elman 神经网络;组合模型;用水量预测

中图分类号: TV213.4

文献标志码: A

文章编号: 1000-7709(2023)12-0028-04

1 引言

水资源是社会经济发展的基础资源之一,对城市的生存与发展至关重要。然而,随着经济的迅猛发展,城镇化进程加快,城市人口持续增长,导致城市用水需求剧增,对城市供水调度工作提出了较大的挑战。对城市用水现状进行科学分析并准确预测用水总量,掌握城市对未来用水量的需求,有利于处理水资源供需之间的平衡,对水资源的开发利用和城市用水系统的改进具有重要意义^[1]。城市用水量受城市水资源总量、气候、人口、经济等诸多不确定因素的影响^[2],用水量与影响因素之间关系复杂,而机器学习算法可以较好地处理该非线性关系,因此被广泛应用在城市用水量的预测中,常用模型有人工神经网络、灰色预测、长短期记忆网络、支持向量机及一些组合模型等^[3],均具有良好的预测效果。南京市位于长江下游中部地区,全市多年平均本地水资源可利用总量为 $17.93 \times 10^8 \text{ m}^3$,据统计南京市 2019 年的总用水量为 $36.33 \times 10^8 \text{ m}^3$,用水需求的增长使水资源成为影响城市经济可持续发展的重要因素,

用水量的准确预测对水资源的长远规划和配置具有重要意义。因此,在分析南京市用水量现状基础上,本文依托用水量现状的变化趋势,结合灰色预测和 Elman 神经网络的优势组成灰色 Elman 神经网络,预测南京市各区用水量及全市总用水量,旨在有效控制区域用水量,实现“四水四定”。

2 研究方法

2.1 Elman 神经网络模型

Elman 神经网络的结构一般分为输入层、隐含层、输出层和承接层四层,每层神经元之间通过权值和阈值连接,通过对权值和阈值的调整优化网络性能。与传统的神经网络相比,Elman 神经网络特殊的承接层可用于记忆隐含层前一刻的输出值,增强了网络的内部反馈机制,使网络处理动态信息的能力提高,具有适应时间变化特征的能力^[4-5]。Elman 神经网络数学模型为:

$$\begin{cases} \mathbf{x}_c(k) = \mathbf{y}_1(k-1) \\ \mathbf{y}_1(k) = f(\mathbf{w}_1 \mathbf{x}_1(k-1) + \mathbf{w}_2 \mathbf{x}_c(k)) \\ \mathbf{y}(k) = g(\mathbf{w}_3 \mathbf{y}_1(k)) \end{cases} \quad (1)$$

式中, $\mathbf{y}_1(k)$ 为隐含层的输出向量; $\mathbf{x}_c(k)$ 为承接

收稿日期: 2023-02-27,修回日期: 2023-03-27

基金项目: 国家重点研发计划(2018YFC1506606);江苏省碳达峰碳中和科技创新专项资金项目(BK20220017)

作者简介: 王浏琳(2000-),女,硕士研究生,研究方向为应用气象学,E-mail:1830373339@qq.com

通讯作者: 张方敏(1983-),女,博士、教授,研究方向为应用气象学,E-mail:fmin.zhang@nuist.edu.cn

层的输出向量; $f(\cdot)$ 为隐含层节点传递函数, 本文选用 tansig 函数; $x_1(k-1)$ 为网络的输入向量; ω_1 为输入层到隐含层的权值向量; ω_2 为承接层到隐含层的权值向量; $y(k)$ 为最终输出向量; w_3 为隐含层到输出层的权值向量; $g(\cdot)$ 为输出层线性函数, 在网络中起线性加权作用。

2.2 灰色预测模型

灰色预测 GM(1,1)模型是灰色系统中最常用的一种动态模型, 通过推断杂乱的小样本数据序列, 并进行规律性地表达, 实现对未来变量的定量预测^[6]。使用灰色 GM(1,1)模型对南京市用水量进行建模预测前, 先要对用水量原始数列的级比进行检验, 对其是否可以建立灰色预测 GM(1,1)模型进行判断。若用来建模的原始数列的级比 $\lambda(k)$ 均落在可容覆盖区间 $X = (e^{-\frac{2}{n+1}}, e^{\frac{2}{n+1}})$ 内, 则原始数列 $x^{(0)}$ 可以建立灰色 GM(1,1)模型并进行灰色预测。

针对每个城区的用水量构建原始数据序列 $x^{(0)}$, 设非负原始数列 $x^{(0)}$ 为:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)) \quad (2)$$

原始数列的级比 $\lambda(k)$ 计算公式为:

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)} \quad k=2, 3, \dots, n \quad (3)$$

对经过可行性检验的原始数列 $x^{(0)}$ 进行累加计算生成数列:

$$\begin{cases} x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \\ x^{(1)}(t) = \sum_{i=1}^t x^{(0)}(i) \quad t=1, 2, \dots, n \end{cases} \quad (4)$$

对一次累加数列 $x^{(1)}$ 建立 GM(1,1)模型的时间响应方程:

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = \mu \quad (5)$$

式中, a 为发展系数, 反映控制系统发展态势的大小; μ 为灰作用量, 反映数据的变化关系。

利用最小二乘法求 a, μ , 即:

$$(\alpha, \mu)^T = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}^N \quad (6)$$

其中 $\mathbf{B} = \begin{bmatrix} -\frac{1}{2(X^1(1)+X^1(2))} & 1 \\ \vdots & \vdots \\ -\frac{1}{2(X^1(n-1)+X^1(n))} & 1 \end{bmatrix} \quad (7)$

$$\mathbf{Y}^N = \begin{bmatrix} x^0(2) \\ x^0(3) \\ \vdots \\ x^0(n) \end{bmatrix} \quad (8)$$

式中, \mathbf{B} 为数据矩阵; \mathbf{Y}^N 为数据向量。

解方程得出 $x^{(1)}$ 的预测值, 然后经累减还原得到原始数列 $x^{(0)}$ 的预测值 $\hat{x}^{(0)}$:

$$\begin{cases} \hat{x}^{(1)}(t+1) = (x^0(1) - \mu/a) e^{-at} + \mu/a & t=0, 1, \dots, n \\ \hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t) & t=1, 2, \dots, n \end{cases} \quad (9)$$

2.3 基于遗传算法的灰色 Elman 神经网络模型

灰色预测 GM(1,1)模型可以使用少量的原始数据完成预测, 但对波动较大的数据拟合能力差, 而 Elman 神经网络可以很好地处理非线性问题, 利用两者的组合模型进行预测可以很好地发挥各自的优势。

Elman 神经网络的初始参数一般随机选取, 本文选用遗传算法 GA 进行全局搜索, 对网络的权值进行优化, 提高网络性能。先用灰色预测 GM(1,1)模型对用水量的原始序列 $X^{(0)}$ 进行预测, 得到预测值 $\hat{X}_i^{(0)}$, 然后将 GM(1,1)模型的预测值和用水量的影响因素作为输入变量, 输入到优化好的 Elman 神经网络中进行用水量预测, 最后得到预测结果。基于遗传算法优化的灰色 Elman 神经网络模型的流程见图 1, 具体步骤如下。

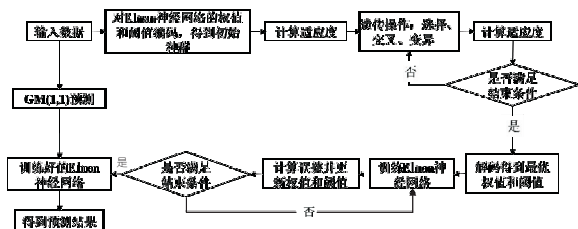


图 1 灰色 Elman 神经网络流程图

Fig. 1 The flow chart of GM-Elman neural network

步骤 1 确定 Elman 神经网络的结构, 主要包括隐藏层神经元的数量和其他网络参数。

步骤 2 种群初始化。设置进化次数、种群规模、交叉概率和变异概率, 对 Elman 神经网络的权值进行实数编码, 随机生成初始种群。

步骤 3 计算适应度。对初始种群计算一次, 然后对经过选择、交叉和变异后产生的每代新种群都计算一次。在网络训练过程中, 输出值和期望值之间的误差越小模型越优, 以实际输出值与期望值之间的误差平方和的倒数作为适应度 F , 计算公式为:

$$F_i = \frac{1}{\sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad i=0, 1, \dots, n \quad (10)$$

式中, x_i, \hat{x}_i 分别为期望输出及实际输出。

步骤 4 循环进行选择、交叉、变异和适应度

计算等操作步骤,形成新一代,直到达到设置的最大进化次数,结束迭代。

步骤 5 对适应度最大的个体进行解码,从而得到网络的最优权值,将最优权值传递给 Elman 神经网络,对网络进行训练,更新权值并计算误差。

步骤 6 将用水量的原始数据输入灰色预测 GM(1,1)模型中进行预测,再将灰色模型的预测结果和影响因素数据输入到训练好的 Elman 神经网络模型中进行预测,得到最终的用水量预测值。

3 南京市用水量现状分析及用水量预测

3.1 南京市用水现状分析

根据南京市水资源公报,南京市的用水总量主要由生活用水、工业用水、农业用水、生态环境用水和林牧渔畜用水 5 部分组成,2009~2019 年的年平均用水量为 $42.02 \times 10^8 \text{ m}^3$,其中工业用水和农业用水占比较高,两者在 2010 年占比最高可达到 73.1%;林牧渔畜用水和生态环境用水占用水总量的比重较低,最低年份在 2011 年,占比 6.25%(图 2)。南京市用水总量在 2009~2017 年的变化较为平稳,2017~2018 年用水总量突然急剧降低,调查发现,国家在此期间实行了水资源消耗总量和强度双控制,导致用水总量急剧减少。其中,生活用水与用水总量的变化趋势几乎相同;农业用水在 2011 年消耗量最大为 $18.61 \times 10^8 \text{ m}^3$,在 2017~2018 年也急剧降低;工业用水在 2010 年消耗量最大为 $18.61 \times 10^8 \text{ m}^3$,在 2017~2018 年虽没有急剧减少,但较之前年份有所降低,整体上与农业用水呈相反趋势波动,两者共同作用影响用水总量的变化。

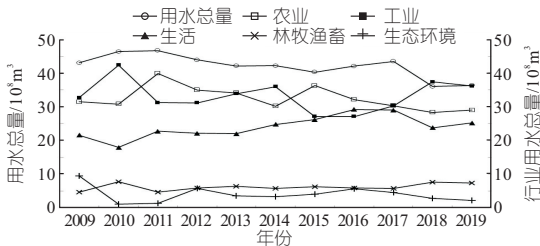


图 2 南京市全市用水量组成及变化

Fig. 2 Change of water consumption in Nanjing

3.2 南京市用水量预测

3.2.1 数据来源和预处理

研究数据来源于南京市统计年鉴、南京市气象观测站、南京市水资源公报、南京市国民经济和社会发展统计公报等 2009~2019 年的数据。基

于全国及南京市水资源动态变化驱动因素的研究结果,选取 8 个因子作为基于灰色 Elman 神经网络算法的南京市用水量预测模型的输入项,即供水总量、人口密度、城镇化率、农林牧渔业增加值、工业增加值、绿化覆盖率、人均 GDP 和年降水量^[3]。

对于南京市各区,将 2009~2018 年的数据作为训练集,2019 年的数据作为验证集;模型网络输入层节点为 8 个,对应影响用水量的 8 个因子;输出层节点为 1 个,对应用水量;同时,设置网络最大训练步数为 10 000,利用该模型对南京市各区用水量及全市总用水量进行预测。为了消除各影响因素之间因为量纲差异产生的影响,在数据输入模型进行训练前,先对所有数据进行归一化处理,即:

$$X_{\text{norm}} = (X - X_{\text{min}}) / (X_{\text{max}} - X_{\text{min}}) \quad (11)$$

式中, X_{norm} 为处理后的数据; X 为处理前的数据; X_{max} 、 X_{min} 分别为数据序列中的最大值和最小值。

3.2.2 基于灰色 Elman 神经网络的用水量预测

为便于比较预测结果,选用相对误差 δ 作为模型预测结果的误差评价指标,相对误差可以反映结果的可靠程度,其计算公式为:

$$\delta = [(y - \hat{y}) / y] \times 100\% \quad (12)$$

式中, y 为实测值; \hat{y} 为预测值。

使用灰色 Elman 神经网络预测的南京市全市总用水量结果见表 1。由表 1 可知,基于灰色 Elman 神经网络预测的 2009~2019 年南京市总用水量结果的变化趋势与真实值基本一致,多年平均相对误差为 1.55%,其中 2013~2015 年的预测结果相对误差较大,预测误差最大的年份在 2013 年为 3.41%。总体来说,灰色 Elman 神经网络模型预测值相对误差均小于 3.5%,且在较多年份的总用水量预测中相对误差低于 1.5%,预测效果良好。2017 年以后,在国家节水政策的管理下,南京市生活用水和农业用水陡然下降造成城市总用水量减少,但经济社会的快速发展和人口的急速增加,使工业、生活用水不会持续下降,南京市未来几年用水量应该比较稳定。

表 1 南京市总用水量预测结果对比

Tab. 1 Comparison of prediction results of total water consumption in Nanjing

年份	实测值 / 10^8 m^3	预测值 / 10^8 m^3	相对误差 /%	年份	实测值 / 10^8 m^3	预测值 / 10^8 m^3	相对误差 /%
2009	43.01	43.04	-0.07	2015	40.24	39.23	2.51
2010	46.36	45.78	1.25	2016	42.02	41.85	0.39
2011	46.64	45.32	2.83	2017	43.46	44.06	-1.37
2012	44.85	44.01	-0.36	2018	36.04	35.97	0.18
2013	42.11	43.55	-3.41	2019	36.33	35.47	2.37
2014	42.14	43.11	-2.31				

表 2 为使用灰色 Elman 神经网络对 2019 年南京市各区用水量的预测结果。由于各区间用水总量的组成占比不同,用水量影响因子亦存在一定差异,导致各区之间使用同一用水量预测模型的预测结果的相对误差差异较大。由表 2 可看出,秦淮区的预测误差最大,相对误差为 8.2%;六合区的预测效果最好,预测值相对误差为 0.32%;2019 年南京市总用水量的相对误差为 2.37%。

表 2 2019 年南京市各区用水量预测结果对比

Tab. 2 Comparison of prediction results of water consumption in each district of Nanjing in 2019

行政 分区	测值 / 10^8 m^3	预测值 / 10^8 m^3	相对 误差/%	行政 分区	测值 / 10^8 m^3	预测值 / 10^8 m^3	相对 误差/%
江北新区	6.63	6.29	5.17	雨花台	1.91	1.83	4.14
玄武	1.08	1.15	-6.81	浦口	1.53	1.42	7.30
秦淮	0.73	0.79	-8.20	江宁	6.64	6.61	0.41
建邺	0.49	0.46	6.99	六合	3.51	3.50	0.32
鼓楼	1.00	0.95	4.55	溧水	2.88	2.85	1.01
栖霞	7.77	7.35	5.40	高淳	2.16	1.98	8.21

根据已有城市用水量预测的研究成果可知,基于机器学习算法的模型在城市用水量预测中取得较好结果^[3,7,8]。本文使用灰色预测 GM(1,1)模型与 Elman 神经网络组合模型预测南京市全市总用水量,预测效果亦较好,相对误差小于 3.5%。表明该模型可准确预测南京市用水量,有利于实现用水总量的控制,为南京地区水资源的高效管理和利用提供科学依据。

4 结论

a. 南京市 2009~2019 年的年平均用水量为 $42.02 \times 10^8 \text{ m}^3$,其中工业、农业用水占总用水量的比重较高,对总用水量的变化起着至关重要的作用。

b. 在南京市 2009~2019 年总用水量预测中,灰色预测 GM(1,1)模型与 Elman 神经网络组合

模型预测结果的相对误差均小于 3.5%,预测结果的多年平均相对误差为 1.55%。对南京市各区用水量使用灰色预测 GM(1,1)模型与 Elman 神经网络组合模型进行预测,各区之间预测效果差异较大,但预测结果的相对误差均小于 10%,预测效果较好。

c. 本文选用了 8 个影响因素作为模型输入,预测南京市用水量,模型对全市总用水量的预测效果良好,故可用于近期南京市用水总量的预测,具有一定参考价值。

参考文献:

- [1] 杨舒媛,王军,魏保义,等.城市需水预测方法综述及应用调研[J].北京规划建设,2016(4):66-68.
- [2] 张真真,单延功,陈家栋,等.南京市水资源脆弱性动态变化及驱动因素分析[J].水利发展研究,2022,22(3):62-68.
- [3] 陈嘉彤,温立书,谭雅心.基于灰色预测和 Elman 神经网络的全国用水量预测[J].江西科学,2018,36(6):961-967.
- [4] 刘拥民,罗皓懿,胡珊.基于 Elman 神经网络的 COVID-19 传播特性预测[J].计算机应用与软件,2022,39(10):42-48,140.
- [5] 张子阳,孙彦广.基于灰色 Elman 神经网络转炉吹氧量的预测[J].计算机应用与软件,2018,35(11):103-107.
- [6] 苏琪,王海波,施晓辰,等.基于灰色预测模型的参数寻优方法及能源预测应用[J].南昌大学学报(理科版),2022,46(3):371-378.
- [7] 白鹏,龙秋波.3 种用水量预测方法在京津冀地区的适用性比较[J].水资源保护,2021,37(2):102-107.
- [8] 迟道才,唐延芳,顾拓,等.灌溉用水量的并联型灰色神经网络预测[J].农业工程学报,2009,25(5):26-29.

Prediction of Nanjing Water Consumption Based on Neural Network

WANG Liu-lin^{1a,1b}, CHEN Jia-dong², ZHANG Fang-min^{1a,1b}

(1a. Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters; 1b. Jiangsu Key Laboratory of Agricultural Meteorology, College of Applied Meteorology, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2. Nanjing Branch of Jiangsu Hydrological and Water Resources Survey Bureau, Nanjing 210008, China)

Abstract: It is of great significance for the long-term planning and allocation of Nanjing water resources to analyze the current situation of Nanjing water consumption, make reasonable prediction of Nanjing water consumption and master the future water demand. The analysis of water consumption in Nanjing from 2009 to 2019 showed that industrial and agricultural water consumption accounted for a large proportion of total water consumption in Nanjing, which played a crucial role in the change of total water consumption. The combined model of grey GM(1,1) model and Elman neural network was used to forecast the water consumption of all districts and the total water consumption of Nanjing. The results show that the grey Elman neural network model has a good prediction effect on the total water consumption of Nanjing City from 2009 to 2019. The relative errors of the forecasts were all less than 3.5%, and the average relative errors of the predicted results over the years were 1.55%; The relative error of the forecast results is less than 8.5% in the forecast of the water consumption of all districts in Nanjing in 2019. The model used in this paper can accurately predict the water consumption of Nanjing, which is of great significance to effectively control the regional water consumption and realize the principle of "four water and four determinations".

Key words: machine learning; grey Elman neural network; combined model; water consumption prediction