

基于 KS-PSO-BP 神经网络耦合模型的需水预测

——以江苏省为例

李 恩¹, 孙伯明², 孙晓文³, 赵 敏², 姚向阳⁴, 颜 冰¹

(1.南京水利科学研究院 水文水资源研究所, 南京 210029; 2.江苏省水利工程科技咨询股份有限公司资源环境部, 南京 210029; 3.江苏省水资源服务中心 水资源规划科, 南京 210028; 4.张家港市水务局 长江防洪工程管理处, 江苏 苏州 215621)

摘 要:需水预测是区域水资源开发利用中的关键环节, 鉴于江苏省需水影响因素涉及范围广、样本系列值少、需水波动幅度大等特点, 提出了一种粒子群优化与正态区间估计联合优化的 KS-PSO-BP 耦合神经网络模型, 模型使用主成分分析法与灰色关联分析法筛选需水影响因素, 加入粒子群优化算法优化 BP 神经网络, 通过引入 K-S 正态检验与正态区间估计, 构建 KS-PSO-BP 神经网络需水预测模型, 并对江苏省需水量进行模拟预测。结果表明: PSO-BP 神经网络在预测结果相对误差上优于 BP 神经网络, 对江苏省 2021—2023 年进行需水预测时, PSO-BP 神经网络预测相对误差区间分别为 $-1.55\% \sim 1.14\%$ 、 $-1.65\% \sim 1.23\%$ 、 $-1.64\% \sim 1.18\%$, 基于 KS-PSO-BP 耦合神经网络模型预测相对误差区间分别为 $-0.23\% \sim 0.09\%$ 、 $-0.28\% \sim 0.01\%$ 、 $-0.28\% \sim 0.02\%$, 可见基于 KS-PSO-BP 耦合神经网络模型极大缩小了误差区间, 模型预测结果更稳定、更准确, 与江苏省实际用水状况更接近, 该模型可作为江苏省需水预测的一种方法。

关键词:需水预测; 主成分分析; 灰色关联分析; PSO-BP 神经网络; K-S 正态检验

中图分类号: TV213.4

文献标志码: A

文章编号: 1001-5485(2026)07-0065-07

Water Demand Prediction Based on KS-PSO-BP Neural Network Coupled Model: A Case Study of Jiangsu Province

LI En¹, SUN Bo-ming², SUN Xiao-wen³, ZHAO Min², YAO Xiang-yang⁴, YAN Bing¹

(1. Department of Hydrology and Water Resources, Nanjing Hydraulic Research Institute, Nanjing 210029, China; 2. Resources and Environment Department of Jiangsu Province Water Engineering Sci-tech Consulting Co., Ltd., Nanjing 210029, China; 3. Department of Water Resources Planning, Jiangsu Provincial Water Resources Service Center, Nanjing 210028, China; 4. Department of Yangtze River Flood Control Management, Zhangjiagang Water Authority, Suzhou 215621, China)

Abstract: [Objective] Current research on water demand prediction primarily focuses on improving prediction accuracy, while studies on interval prediction of water demand remain limited. This study aims to further improve prediction accuracy, reduce error interval and reflect the actual regional water demand by incorporating the Kolmogorov-Smirnov (K-S) normal interval estimation into BP (back propagation) neural network model. [Methods] In view of a wide range of water demand influencing factors, limited sample series, and significant demand fluctuations in Jiangsu Province, we coupled the K-S normal interval estimation with particle swarm optimization (PSO) and BP neural network. Specific processes are as follows: the degree of influence and the number of key factors were preliminarily identified through principal component analysis (PCA), followed by the calculation of grey relational grades (GRG) to determine the final water demand influencing factor index. This index was utilized as input for both BP and PSO-BP neural networks. The final model was then selected by comparing performance indicators, including relative error, Nash-Sutcliffe efficiency (NSE), and coefficient of determination (R^2). Using multiple fitting iterations of the PSO-BP neural network, the sample size was expanded, and the K-S normality test and normal interval estimation were performed to further narrow the relative error intervals of the water demand prediction re-

收稿日期: 2025-04-26; 修回日期: 2025-07-23; 接受日期: 2025-07-29

基金项目: 江苏省水利科技项目(2024024); 中央级公益性科研院所基本科研业务费专项资金项目(Y525005, Y525007)

作者简介: 李 恩(1999-), 男, 安徽宿州人, 工程师, 博士研究生, 主要研究方向为水文学及水资源。E-mail: eli@nhri.cn

通信作者: 颜 冰(1992-), 男, 安徽滁州人, 高级工程师, 博士, 研究方向为水文学及水资源。E-mail: byan@nhri.cn

sults. [**Results**] 1) Eight primary influencing factors of water demand were identified for Jiangsu Province: population, GDP, added value of secondary industry, added value of tertiary industry, per capita urban domestic water consumption, per capita rural domestic water consumption, irrigation water quota per mu ($1 \text{ mu} \approx 666.7 \text{ m}^2$), and water consumption per 10 000 yuan of industrial added value. 2) Both BP and PSO-BP neural networks exhibited strong performance in fitting historical water demand data, demonstrating their feasibility for future water demand prediction. Specifically, the PSO-BP neural network outperformed the standard BP network in the relative error of water demand prediction. The maximum relative errors for the training, validation, and testing samples were 2.23%, 0.88%, and 1.19%, respectively, with an average training error of 1.03%. The NSE and R^2 reached 0.98 and 0.99, respectively. 3) The selection of sample size significantly influenced the prediction results of water demand. With the increase of the sample size, the average value of samples was closer to the real value, and the overall prediction results of samples were more stable. The increase of sample size would reduce the benefit of improving the accuracy of the overall prediction results. 4) The K-S normality test and normal interval estimation stabilized the prediction results, substantially narrowed the error intervals, and better reflected the actual water demand. At a 95% confidence level, the relative error intervals of water demand prediction in 2021, 2022, and 2023 in Jiangsu Province were reduced to $-0.23\% - 0.09\%$, $-0.28\% - 0.01\%$, and $-0.28\% - 0.02\%$, respectively. [**Conclusion**] The KS-PSO-BP neural network coupled model significantly reduces the error range and provides more stable and accurate prediction results that closely align with actual water demand in Jiangsu Province. The model serves as an effective method for regional water demand prediction and offers valuable guidance for future water resources planning in Jiangsu Province.

Key words: water demand prediction; principal component analysis; grey relational analysis; PSO-BP neural network; K-S normality test

0 引 言

区域需水预测是水资源规划与管理的重要内容,随着城镇化进程的加快以及人民生活水平的不断提高,区域水资源短缺和需水量日益增长的矛盾更加突出,通过识别需水预测主要影响因子,预测未来需水规模,对于科学制定水资源开发利用方案具有重要意义^[1-4]。受人口、经济结构等不同因素影响,区域需水量与各影响因素之间通常呈现非线性关系,具有较强的波动性和不确定性。为提高区域需水量预测精度,国内外学者开展了大量研究,形成了以传统预测和现代预测为主流的需水预测方法,其中传统预测以趋势法、多元线性回归法等为代表,现代预测以反向传播(Back Propagation, BP)神经网络、灰色预测模型等为代表。Bata 等^[5]基于神经网络模型预测了不同时段需水量。Stelzl 等^[6]采用多元线性回归模型预测不同水平年峰值需水量。李嘉欣等^[7]运用 Holt 线性趋势法、灰色预测模型、定额法和多元线性回归方法对黄河流域 2035 年需水量进行预测,结果表明多元线性回归模型预测效果最好。毛青等^[8]基于趋势法、多元线性回归和 BP 神经网络,利用相对误差-反距离

权重法构建了一种组合模型,提高了预测精度。杨登元等^[9]采用改进主成分分析(Principal Component Analysis, PCA)-BP 神经网络模型预测海宁市规划年需水量,并与传统定额法的需水预测结果进行对比分析,结果表明改进 PCA-BP 神经网络模型预测结果更加真实。由于 BP 神经网络非线性拟合能力效果好,近年来,该模型逐步应用到区域需水预测中,但其收敛速度慢、易陷入局部最优的特性并未改变,在需水预测中,受到各类因素相互交叉影响,导致预测结果偏离真实值,单一采用 BP 神经网络模型预测区域需水量难以得到理想结果^[10-12]。为加强 BP 神经网络最优组合的搜索能力,最小化模型输出误差,通过粒子群算法优化 BP 神经网络,同时为反映需水量本身具备的波动性特点,引入 K-S 正态检验与正态区间估计,构建粒子群算法(Particle Swarm Optimization, PSO)与正态区间估计联合优化的 KS-PSO-BP 神经网络模型,实现预测结果的精确化、稳定化^[13-15]。

本文以江苏省为例,采用 PSO-BP 神经网络对江苏省历史数据进行模拟训练,并引入 K-S 正态检验与正态区间估计,构建 KS-PSO-BP 神经网络耦合模型,以求实现提升预测精度、缩小模拟误差区间、反映江苏省实际需水情况。

1 研究区概况

江苏省作为沿海地区最具经济活力的省份之一,2023年地区生产总值达12.8万亿元,第二、第三产业为支柱型产业,全年三次产业结构比例为4:44.4:51.6。为适应新时期水利发展要求,江苏省在水资源利用方面积极进行“控总量、提效率”,并取得了良好成绩,2023年全省用水总量571.4亿 m^3 ,控制在国家下达的用水总量指标内。按产业结构分,第一产业用水266.9亿 m^3 ,占用水总量的46.7%,第二产业用水282.8亿 m^3 ,占用水总量的49.5%,第三产业用水21.7亿 m^3 ,占用水总量的3.8%。2004—2023年,江苏省用水结构并无明显变化,第一、第二产业用水仍占据主导地位;万元GDP用水量31.0 m^3 ,万元工业增加值用水量17.7 m^3 ,分别比2020年下降19.3%和13.3%,保持全国领先水平;耕地实际灌溉亩均用水量400.2 m^3 ,城镇居民人均生活用水量147.5 L/d ,均低于全国平均水平。江苏省是南水北调东线工程起点,既担负着对外供水的任务,又要确保省内用水安全,随着全省产业变革的深入推进以及人民生活水平的提高,区域经济平稳上升,水资源需求量进一步扩大。

2 数据来源与研究方法

2.1 数据来源

本研究使用的人口、GDP、第一产业增加值、第二产业增加值、第三产业增加值、农作物播种面积等数据来源于2005—2024年《江苏省统计年鉴》,用水总量、万元GDP用水量、万元工业增加值用水量、降雨量、城镇生活人均用水量、农村生活人均用水量、亩均灌溉用水量、工业用水重复利用率等数据来源于2004—2023年《江苏省水资源公报》。

2.2 研究方法

2.2.1 主成分分析法

主成分分析法是一种利用降维思想,将多指标问题转化为少数综合指标的简化数据集技术,便于消除指标之间的相关影响^[16-18],计算步骤如下:

对指标进行标准化处理,形成数据矩阵。

$$\mathbf{X}_{mn} = (X_1, X_2, \dots, X_m) \quad (1)$$

式中: $\mathbf{X}_{mn}$ 为数据矩阵; $X_1, X_2, \dots, X_m$ 分别为第1,第2, ..., 第 m 个指标,每个指标有 n 个样本。

对 $\mathbf{X}_{mn}$ 进行线性变换。

$$\begin{cases} \mathbf{Y}_{kn} = \mathbf{P}_{km} \mathbf{X}_{mn} & ; \\ \mathbf{Y}^T = (y_1, y_2, \dots, y_k) & (k < m) \end{cases} \quad (2)$$

式中: $\mathbf{Y}_{kn}$ 为变换后的数据矩阵; $\mathbf{P}_{km}$ 为特征向量组成的矩阵; $\mathbf{Y}^T$ 为变换后数据矩阵的转置, k 为主成分, m 为原始指标, $y_1, y_2, \dots, y_k$ 分别为第1,第2, ..., 第 m 个指标,每个指标有 n 个样本。

将变换后数据矩阵的协方差矩阵作为对角矩阵,即

$$\text{Var}(\mathbf{Y}) = \frac{1}{n} \mathbf{Y} \mathbf{Y}^T = \mathbf{P} \text{Var}(\mathbf{X}) \mathbf{P}^T \quad (3)$$

式中: $\text{Var}(\mathbf{Y})$ 为对角矩阵; $\mathbf{P}$ 为特征向量矩阵; $\mathbf{P}^T$ 为 $\mathbf{P}$ 的转置; $\text{Var}(\mathbf{X})$ 为协方差矩阵。

2.2.2 灰色关联分析法

灰色关联分析法能够定量描述各指标对选定指标的影响程度^[19-20],计算过程如下。

设定相关指标:

$$\begin{cases} X_0(k) = [X_0(1), X_0(2), \dots, X_0(m)] & ; \\ X_i(k) = [X_i(1), X_i(2), \dots, X_i(n)] & \end{cases} \quad (4)$$

式中: $X_0(k)$ 为选定指标; $X_i(k)$ 为影响因素指标, $i=1, 2, \dots, n$ 。

消除量纲影响:

$$X'_i = X_i / X_i(1) \quad (5)$$

计算绝对差值:

$$\Delta_i(k) = |X'_0(k) - X'_i(k)| \quad (6)$$

式中: $\Delta_i(k)$ 为第 i 个自变量与因变量在第 k 个样本点上的绝对差值; $X'_0(k)$ 为目标指标在第 k 个样本的值; $X'_i(k)$ 为影响因素在第 k 个样本的值。

计算关联系数:

$$\begin{cases} \Delta_{\max} = \max_i \max_k |X'_0(k) - X'_i(k)| & ; \\ \Delta_{\min} = \min_i \min_k |X'_0(k) - X'_i(k)| & ; \\ \xi_i(k) = (\Delta_{\min} + \rho \Delta_{\max}) / (\Delta_i(k) + \rho \Delta_{\max}) & \end{cases} \quad (7)$$

式中: ρ 为分辨系数; ρ 越小关联系数差异越显著; $\Delta_{\max}$ 为绝对差值的最大值; $\Delta_{\min}$ 为绝对差值的最小值。

计算关联度 r_i ,关联度越大,相关性越强,即

$$r_i = \frac{1}{n} \sum_k \xi_i(k) \quad (8)$$

式中 $\xi_i(k)$ 为关联系数。

2.2.3 BP神经网络

BP神经网络基于输入信号正向传播,由误差反向传播算法训练多层前馈网络。其基本结构包括输入层、隐藏层和输出层,通过模拟脑神经信号,多次训练样本值,根据误差反向传播方式与规律进行不断调整,确定影响因素与结果的关系^[21],基本原理如下:

$$H_j = f\left(\sum \omega_{ij} x_i + b_j\right) \quad (9)$$

式中: H_j 为隐藏层; f 为激活函数; ω_{ij} 为权重; x_i 为输入层; b_j 为阈值。

$$\begin{cases} E(\omega, b) = \frac{1}{n} \left[\sum (Y^n - y^n)^2 \right] \\ \omega_{ij} = -\eta \frac{\partial E(\omega, b)}{\partial \omega} \end{cases} \quad (10)$$

式中: $E(\omega, b)$ 为误差函数; Y^n 为实际值; y^n 为模拟值; η 为学习率; ∂ 为偏微分符号。

由于本文采用 20 a 时序的小样本数据作为 BP 神经网络输入, 为有效构建和应用 BP 神经网络, 首先要对样本质量进行优化, 检查历年用水量分布情况, 剔除异常值, 通过提取主要影响因素, 确保需水影响类别分布均匀。其次隐藏层节点数的设置需遵循小样本适配原则, 既要避免欠拟合又要防止过拟合, 通过数据“质”的提升弥补“量”的不足, 同时使模型复杂度与数据规模相匹配, 节点数可采用式(11)确定, 即

$$H_0 = \sqrt{X_0 + Y_0} + \alpha \quad (11)$$

式中: H_0 为隐藏层节点数; X_0 为输入层节点数; Y_0 为输出层节点数; α 为调节因子。

2.2.4 PSO-BP 神经网络

PSO-BP 神经网络是基于粒子群算法的 BP 神经网络, 将权重和偏置参数编码为 PSO 粒子群中的粒子, 初始化粒子的位置和速度, 选择或设计一个适应度函数, 评估粒子(神经网络参数集)的好坏, 通过速度和位置的迭代更新, 最终收敛于最优解或近似最优解, 在此过程中, PSO 算法负责搜索这些参数的最优组合, 使得神经网络的输出误差最小化^[22]。

2.2.5 K-S 正态检验与正态区间估计

K-S 正态检验用于检验个体样本属于特定的概率分布, 其基本原理是将样本系列值的累计概率分布与正态分布进行对比分析, 判断样本数据是否符合正态分布^[23]。正态区间估计则是对正态分布 $N(\mu, \sigma^2)$ 中的 μ 进行区间估计, 若总体样本 X 包含 $X_1, X_2, \dots, X_n$, 样本均值为 $\bar{X}$, 样本方差为 S^2 , 则有

$$T = \frac{|\bar{X} - \mu|}{S/\sqrt{n}} \sim t(n-1) \quad (12)$$

式中: T 为检验统计量; μ 为总体均值; n 为原始样本容量; S 为样本标准差。

在置信度为 $1 - \alpha$ 下, μ 的置信区间为

$$\left(\bar{X} - t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{N}}, \bar{X} + t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{N}} \right) \quad (13)$$

式中: $t_{\frac{\alpha}{2}}$ 为分位数; N 为模拟样本容量。

2.2.6 KS-PSO-BP 耦合神经网络模型

KS-PSO-BP 耦合神经网络模型的构建可分为 4 个步骤: ①通过主成分分析初步确定需水因子对需水量的影响程度及对应数量; ②计算各影响因子与用水量的灰色关联系数, 确定最终的需水影响因子指标; ③将需水影响因子指标用于 BP 神经网络与 PSO-BP 神经网络的需水预测, 通过相对误差、纳什系数、决定系数等指标对比两者预测精度, 确定最终预测模型; ④通过 PSO-BP 神经网络的多次拟合, 扩充合适样本数量, 并进行 K-S 正态检验与正态区间估计, 进一步缩小预测结果相对误差区间。模型技术路线如图 1 所示。

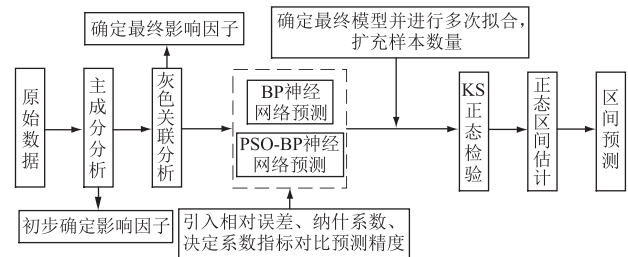


图 1 KS-PSO-BP 耦合神经网络模型构建思路

Fig.1 Framework of KS-PSO-BP coupled neural network model

3 结果与分析

3.1 需水影响因子的率定

以江苏省 2004—2023 年用水量作为因变量, 以人口 (z_1 , 万人)、GDP (z_2 , 亿元)、第一产业增加值 (z_3 , 亿元)、第二产业增加值 (z_4 , 亿元)、第三产业增加值 (z_5 , 亿元)、万元 GDP 用水量 (z_6 , m^3)、降水量 (z_7 , mm)、城镇生活人均用水量 (z_8 , m^3)、农村生活人均用水量 (z_9 , m^3)、农作物播种面积 (z_{10} , hm^2)、亩均灌溉用水量 (z_{11} , m^3)、万元工业增加值用水量 (z_{12} , m^3)、工业用水重复利用率 (z_{13} , %) 为自变量, 进行主成分分析, 见表 1。

表 1 主成分分析结果

Table 1 Results of principal component analysis

主成分	特征值	方差贡献率/%	累计贡献率/%
1	6.929	53.298	53.298
2	3.111	23.935	77.232
3	1.200	9.231	86.464

由表 1 可以看出, 3 个主成分方差累计贡献率达 86.464%, 表明 3 个主成分基本能说明用水量所包含的全部信息, 按照贡献率排序, 第一主成分选定前 7 个指标, 第二主成分选定前 3 个指标, 第三主成分选定前 1 个指标, 提取指标为 $z_1, z_2, z_3, z_4, z_5, z_8, z_9, z_{10}, z_{11}, z_{12}$ 共计 10 个指标, 见表 2。

表2 主成分指标提取

Table 2 Extraction of principal component indicators

第一主成分		第二主成分		第三主成分	
指标	特征向量	指标	特征向量	指标	特征向量
z_3	0.992	z_9	0.931	z_{12}	0.403
z_4	0.991	z_8	0.859		
z_5	0.985	z_{11}	0.717		
z_1	0.980				
z_{10}	0.973				
z_2	0.774				
z_{12}	0.531				

计算所有指标与用水量的灰色关联系数,对照主成分分析结果进行需水影响因子率定,最终确定8个需水影响因子指标,分别为:人口、GDP、第二产业增加值、第三产业增加值、城镇生活人均用水量、农村生活人均用水量、亩均灌溉用水量、万元工业增加值用水量,见表3。

表3 需水影响因子灰色关联系数

Table 3 Grey relational coefficients of water demand influencing factors

指标	关联系数	指标	关联系数	指标	关联系数
z_{11}	0.927	z_{13}	0.874	z_2	0.679
z_{10}	0.921	z_7	0.841	z_6	0.700
z_8	0.920	z_3	0.723	z_5	0.691
z_1	0.905	z_{12}	0.703		
z_9	0.887	z_4	0.695		

3.2 BP神经网络与 PSO-BP神经网络需水预测

根据需水量影响因子率定结果,通过 Python 编写 BP 神经网络、PSO-BP 神经网络代码,以江苏省 2004—2018 年用水量及其影响因子数据为训练样本,分别进行模拟训练,以 2019—2020 年数据为验证样本,选择相对误差(δ)、纳什系数(NSE)、决定系数(R^2)对数据样本进行检验,神经网络模型逐年拟合结果与逐年相对误差见图2。

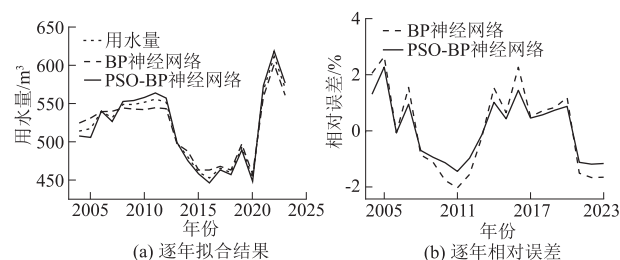


图2 神经网络模型逐年拟合结果与逐年相对误差

Fig.2 Annual fitting results and relative errors of neural network models

为验证 BP 神经网络以及 PSO-BP 神经网络对历史数据的拟合精度,选择训练样本最大相对误差、验证样本最大相对误差、检验样本最大相对误差、训练样本平均相对误差、纳什系数、决定系数对预测结果进行检验。PSO-BP 神经网络各类误差均小于 BP

神经网络,样本纳什系数、决定系数均优于 BP 神经网络,在拟合精度上有明显改观,训练样本最大相对误差、验证样本最大相对误差、检验样本最大相对误差以及训练样本平均相对误差分别为 2.23%、0.88%、1.19%、1.03%,整体样本纳什系数、决定系数分别为 0.98、0.99。2 种方法预测值与真实值之间的平均相对误差均 $<5\%$,纳什系数与决定系数趋近于 1,从侧面印证了预测结果的合理性,同时说明 BP 神经网络以及 PSO-BP 神经网络整体上均可满足需水预测要求。需水预测方法拟合精度统计见表 4。

表4 需水预测方法拟合精度统计

Table 4 Statistics of fitting accuracy of water demand prediction methods

类型	训练样本最大相对误差/%	验证样本最大相对误差/%	检验样本最大相对误差/%	训练样本平均相对误差/%	样本纳什系数	样本决定系数
BP 神经网络	2.66	1.20	1.50	1.38	0.97	0.98
PSO-BP 神经网络	2.23	0.88	1.19	1.03	0.98	0.99

3.3 基于 KS-PSO-BP 耦合神经网络模型需水预测

由于江苏省历史需水系列数据具有波动性和不确定的特点,单一使用 BP 神经网络或者 PSO-BP 神经网络难以真实还原实际需水情况。此外,在进行模拟预测时,各初始条件的变化依旧会造成预测结果的多样性,为尽可能降低这种影响,构建 KS-PSO-BP 神经网络耦合模型,对预测结果进行 K-S 检验,并开展正态区间估计。

(1)根据江苏省 2004—2023 年用水量历史系列值数据,以系列值数据的 75%作为训练样本,即 2004—2018 年,以系列值数据的 10%作为验证样本,即 2019—2020 年,系列值数据的 15%作为对比样本,用于检验预测结果的真实性。通过 PSO-BP 神经网络的多次拟合,挑选训练样本平均相对误差、验证样本最大相对误差 $<2\%$ 的 PSO-BP 神经网络,每年记录 40、60、80、100 个样本,分别统计 2021—2023 年样本频数分布及预测结果,见图 3、表 5。

(2)对所选预测结果进行 K-S 检验,经计算 2021—2023 年样本数据 P 均为 0.2, >0.05 ,表明接受零假设,即样本服从正态分布,通过正态分布均值区间估计,通过 KS-PSO-BP 神经网络耦合模型得到江苏省 2021 年、2022 年、2023 年需水量区间预测结果分别为 566.30 亿 $\sim$ 568.10 亿 m^3 、610.10 亿 $\sim$ 611.89 亿 m^3 、569.80 亿 $\sim$ 571.49 亿 m^3 。江苏省 2021 年、2022 年、2023 年需水量分别为 567.6 亿 m^3 、611.8 亿 m^3 、571.4 亿 m^3 ,结果分析见表 6。

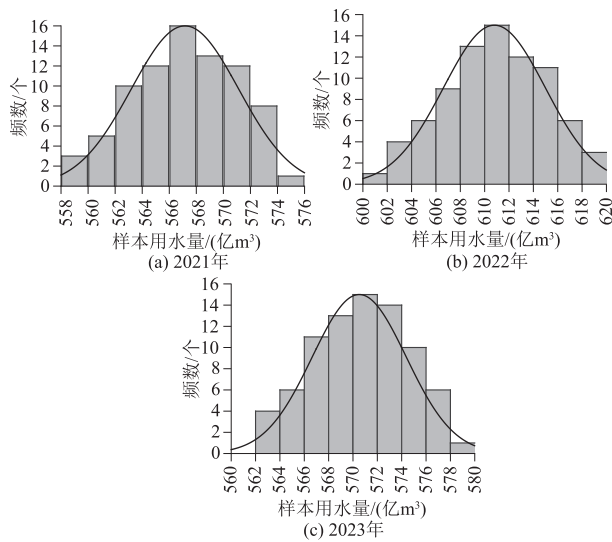


图3 PSO-BP神经网络2021—2023年样本频率分布

Fig.3 Sample frequency distribution of PSO-BP neural network (2021–2023)

表5 PSO-BP神经网络2021—2023年需水量预测结果统计

Table 5 Statistics of water demand prediction results of PSO-BP neural network (2021–2023)

年份	样本数量/个	真实值/ (亿 m ³)	最小模拟值		最大模拟值		平均模拟值/ (亿 m ³)
			数值/ (亿 m ³)	相对误差/%	数值/ (亿 m ³)	相对误差/%	
2021	40	567.6	559.62	-1.41	573.78	1.09	567.02
	60	567.6	559.21	-1.48	573.96	1.12	567.03
	80	567.6	558.80	-1.55	574.07	1.14	567.16
	100	567.6	558.72	-1.56	574.15	1.15	567.07
2022	40	611.8	602.54	-1.51	618.22	1.05	610.64
	60	611.8	602.19	-1.57	618.74	1.13	610.55
	80	611.8	601.71	-1.65	619.33	1.23	610.80
	100	611.8	601.62	-1.66	619.37	1.24	610.82
2023	40	571.4	563.56	-1.37	576.69	0.93	570.01
	60	571.4	563.17	-1.44	577.31	1.03	570.10
	80	571.4	562.03	-1.64	578.14	1.18	570.52
	100	571.4	561.97	-1.65	578.35	1.22	570.44

表6 KS-PSO-BP神经网络耦合模型2021—2023年需水量区间预测结果

Table 6 Interval prediction results from 2021 to 2023 based on KS-PSO-BP neural network coupled model

年份	真实值/ (亿 m ³)	区间预测 (95%置信度)			
		下限/ (亿 m ³)	相对误差/%	上限/ (亿 m ³)	相对误差/%
2021	567.6	566.30	-0.23	568.10	0.09
2022	611.8	610.10	-0.28	611.89	0.01
2023	571.4	569.80	-0.28	571.49	0.02

从表5可知,样本数量的选取对于需水预测结果有一定影响,以2021年为例,当样本数量分别为40、60、80、100个时,样本平均值分别为567.02亿、567.03亿、567.16亿、567.07亿 m³,表明随着样本数量

增加,样本平均值更接近真实值,样本整体预测结果更加稳定,当样本数量由80个变为100个时,样本平均值变化幅度减少,表明增加样本数量对于整体预测结果精度提高的收益降低,因此本次选择 PSO-BP 神经网络中80个样本对应的江苏省2021年、2022年、2023年需水量预测结果,预测相对误差区间分别为-1.55%~1.14%、-1.65%~1.23%、-1.64%~1.18%。

从表6可以看出,在引入K-S正态检验以及正态区间估计后,采用KS-PSO-BP神经网络耦合模型对江苏省2021年、2022年、2023年需水量进行预测时,预测相对误差区间分别为-0.23%~0.09%、-0.28%~0.01%、-0.28%~0.02%。证明KS-PSO-BP神经网络耦合模型可以极大缩小误差区间,使得预测结果更稳定,精确度更高。

4 结 论

本研究结合江苏省需水特性,在 PSO-BP 神经网络的基础上,引入 K-S 正态检验以及正态区间估计,构建 KS-PSO-BP 神经网络耦合模型,对江苏省需水总量进行区间预测,主要结论如下。

(1)BP神经网络以及 PSO-BP 神经网络拟合江苏省历史需水数据表现良好,可用于预测江苏省未来需水量,其中 PSO-BP 神经网络在预测结果相对误差上优于 BP 神经网络,训练样本最大相对误差、验证样本最大相对误差、检验样本最大相对误差以及训练样本平均相对误差分别为2.23%、0.88%、1.19%、1.03%,样本纳什系数、决定系数分别为0.98、0.99。

(2)PSO-BP神经网络预测结果呈现出较强的波动性,对江苏省2021年、2022年、2023年进行需水预测时,预测相对误差区间分别为-1.55%~1.14%、-1.65%~1.23%、-1.64%~1.18%,表现为相对误差区间较大,难以真实还原江苏省实际需水情况。

(3)通过引入K-S正态检验以及正态区间估计后,建立KS-PSO-BP神经网络耦合模型,可以使得预测结果更加稳定,极大缩小相对误差区间,更能反映实际需水情况。在置信度95%的情况下,对江苏省2021年、2022年、2023年进行需水预测时,预测相对误差区间分别为-0.23%~0.09%、-0.28%~0.01%、-0.28%~0.02%。

参考文献 (References):

- [1] 杨皓翔, 梁 川, 崔宁博. 基于加权灰色-马尔可夫链模型的城市需水预测[J]. 长江科学院院报, 2015(7): 15~

- 21.(Yang Hao-xiang,Liang Chuan,Cui Ning-bo.Prediction of Urban Water Demand by Using Weighted Grey-Markov Chain Model[J]. Journal of Changjiang River Scientific Research Institute,2015(7):15-21.(in Chinese))
- [2] 邵 骏,卢满生,杜 涛,等.长江流域水资源生态足迹及其驱动因素[J].长江科学院院报,2021,38(12):19-24,32.(Shao Jun, Lu Man-sheng, Du Tao, *et al.* Water Resources Ecological Footprint in Yangtze River Basin and Its Driving Factors[J]. Journal of Yangtze River Scientific Research Institute, 2021, 38(12): 19-24, 32.(in Chinese))
- [3] 邵 磊,周孝德,杨方廷,等.基于RAGA的GM(1,1)-RBF组合需水预测模型[J].长江科学院院报,2010,27(5):29-33.(Shao Lei, Zhou Xiao-de, Yang Fang-ting, *et al.* Water Demand Prediction Model Constructed by GM(1,1)-RBF Portfolio Neural Network Based on RAGA[J]. Journal of Yangtze River Scientific Research Institute, 2010, 27(5): 29-33.(in Chinese))
- [4] Gabriele A, Biondi D, Gargano R, *et al.* From Deterministic to Probabilistic Forecasts of Water Demand[C]//The 3rd International Joint Conference on Water Distribution Systems Analysis & Computing and Control for the Water Industry (WDSA/CCWI 2024). MDPI, 2024, 69(1):122.
- [5] Bata M H, Carrière R, Ting D S. Short-term Water Demand Forecasting Using Nonlinear Autoregressive Artificial Neural Networks[J]. Journal of Water Resources Planning and Management, 2020, 146(3): 04020008.
- [6] Stelzl A, Pointl M, Fuchs-Hanusch D. Estimating Future Peak Water Demand with a Regression Model Considering Climate Indices[J]. Water, 2021, 13(14): 1912.
- [7] 李嘉欣,彭少明.黄河流域用水现状及需水预测研究[J].人民黄河,2024,46(7):66-71.(Li Jia-xin, Peng Shao-ming. Study on the Current Situation of Water Use and Water Demand Forecast in the Yellow River Basin[J]. Yellow River, 2024, 46(7): 66-71.(in Chinese))
- [8] 毛 青,解阳阳,刘赛艳,等.南水北调受水区工业需水量的多方法组合预测[J].水资源与水工程学报,2024,35(3):51-58,66.(Mao Qing, Xie Yang-yang, Liu Sai-yan, *et al.* Multi-method Combination Prediction of Industrial Water Demand in the Receiving Area of the South to North Water Diversion Project[J]. Journal of Water Resources and Water Engineering, 2024, 35(3): 51-58, 66.(in Chinese))
- [9] 杨登元,鞠茂森,唐德善.基于改进PCA-BP神经网络模型的海宁市需水预测[J].水电能源科学,2024,42(5):68-71,79.(Yang Deng-yuan, Ju Mao-sen, Tang De-shan. Water Demand Prediction in Haining City Based on Improved PCA-BP Neural Network Model[J]. Water Resources and Power, 2024, 42(5): 68-71, 79.(in Chinese))
- [10] Jeyavelu S, Mohan K. A Hierarchical Hybrid Deep-net Framework for Water Demand Forecasting Using OPHO Optimization and Deep Learning Techniques[J]. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2025, 49(5): 5255-5271.
- [11] Jia S, Cui M, Chen L, *et al.* Soybean Water Monitoring and Water Demand Prediction in Arid Region Based on UAV Multispectral Data[J]. Agronomy, 2025, 15(1): 88.
- [12] 董增川,王佳晟,崔 璨,等.基于滚动交叉验证的城市需水预测方法[J].水资源保护,2025,41(3):13-19.(Dong Zeng-chuan, Wang Jia-sheng, Cui Can, *et al.* Method for Urban Water Demand Forecast Based on Rolling Cross-validation[J]. Water Resources Protection, 2025, 41(3): 13-19.(in Chinese))
- [13] Iglesias-Rey A, López Hojas C A, Martínez-Solano F J, *et al.* An Approach Based on the Use of Commercial Codes and Engineering Judgement for the Battle of Water Demand Forecasting[C]//The 3rd International Joint Conference on Water Distribution Systems Analysis & Computing and Control for the Water Industry (WDSA/CCWI 2024). MDPI, 2024: 176.
- [14] Yang M, Gao E, Wang G, *et al.* Medium and Long-term Regional Water Demand Prediction Using Harris Hawks Optimisation-Back Propagation Neural Network Model[J]. Scientific Reports, 2024, 14: 27763.
- [15] 陈伟楠,杨程天.基于主成分分析的PSO-BP模型在城市需水预测中的应用[J].信息与电脑(理论版),2018,30(13):48-50.(Chen Wei-nan, Yang Cheng-tian. Application of PSO-BP Model Based on Principal Component Analysis in Prediction of Urban Water Demand[J]. China Computer & Communication, 2018, 30(13): 48-50.(in Chinese))
- [16] 孔祥仟,陈 园,刘博懿,等.基于主成分分析的神经网络在需水预测中的应用[J].水电能源科学,2018,36(4):26-28.(Kong Xiang-qian, Chen Yuan, Liu Bo-yi, *et al.* Application of Neural Network Based Principal Component Analysis in Prediction of Water Demand[J]. Water Resources and Power, 2018, 36(4): 26-28.(in Chinese))
- [17] 王春娟,冯利华,罗 伟,等.主成分回归在需水预测中的应用[J].水资源与水工程学报,2013,24(1):50-53.(Wang Chun-juan, Feng Li-hua, Luo Wei, *et al.* Application of Principal Component Regression Model to Water Demand Forecast[J]. Journal of Water Resources and Water Engineering, 2013, 24(1): 50-53.(in Chinese))
- [18] 龚 杰,赵起超,娄华超,等.模糊层次分析法在水资源价值评估中的应用:以绵阳市为例[J].长江科学院院报,2022,39(4):34-40.(Gong Jie, Zhao Qi-chao, Lou Hua-chao, *et al.* Application of Fuzzy Analytic Hierarchy Process to Water Resource Value Assessment: Case Study of Mianyang City[J]. Journal of Changjiang River Scientific Research Institute, 2022, 39(4): 34-40.(in Chinese))

- [9] 黄岱. 长江上游宜宾-重庆河段日调节非恒定流传播特性研究[D]. 重庆: 重庆交通大学, 2019. (Huang Dai. Study on the Propagation Characteristics of Daily Regulated Unsteady Flow in Yibin-Chongqing Riverway in the Upper Reach of the Yangtze River[D]. Chongqing: Chongqing Jiaotong University, 2019. (in Chinese))
- [10] 母德伟, 王永强, 李学明, 等. 向家坝日调节非恒定流对下游航运条件影响研究[J]. 四川大学学报(工程科学版), 2014, 46(6): 71-77. (Mu De-wei, Wang Yong-qiang, Li Xue-ming, et al. Study on the Effect of Downstream Navigation by Unsteady Flow Caused by Xiangjiaba Daily Regulation[J]. Journal of Sichuan University (Engineering Science), 2014, 46(6): 71-77. (in Chinese))
- [11] 张有林, 李文杰, 张 谧, 等. 向家坝枢纽非恒定泄流对落锅滩河段航道通航条件的影响[J]. 水运工程, 2018(11): 91-96. (Zhang You-lin, Li Wen-jie, Zhang Mi, et al. Influence of Xiangjiaba Unsteady Flow on Waterway Navigation Conditions in Luogutan Reach[J]. Port & Waterway Engineering, 2018(11): 91-96. (in Chinese))
- [12] 杜泽东, 董先勇, 秦蕾蕾, 等. 溪洛渡水电站坝下游河道非恒定流特性研究[J]. 人民长江, 2023, 54(1): 200-205. (Du Ze-dong, Dong Xian-yong, Qin Lei-lei, et al. Study on Unsteady Flow Characteristics in Downstream Channel of Xiluodu Hydropower Station[J]. Yangtze River, 2023, 54(1): 200-205. (in Chinese))
- [13] 钱红露, 尹维清, 程 稳, 等. 高山峡谷河段枢纽运行对下游水位变幅影响研究: 以金沙江乌东德水电站为例[J]. 人民长江, 2021, 52(4): 100-106. (Qian Hong-lu, Yin Wei-qing, Cheng Wen, et al. Study on Influence of Operation of Hydropower Stations in High Mountain and Canyon River Section on Downstream Water Level Variation: Case of Wudongde Hydropower Station on Jinsha River[J]. Yangtze River, 2021, 52(4): 100-106. (in Chinese))
- [14] 卢程伟, 陈莫非, 张余龙, 等. 断波在朱沱—三峡坝址库区河段传播规律分析[J]. 长江科学院院报, 2021, 38(8): 14-18, 24. (Lu Cheng-wei, Chen Mo-fei, Zhang Yu-long, et al. Propagation of Surge Wave in Three Gorges Reservoir Area Reach from Zhutuo to Dam Site[J]. Journal of Yangtze River Scientific Research Institute, 2021, 38(8): 14-18, 24. (in Chinese))
- [15] 孙 干, 孙然好, 徐火清, 等. 干热河谷梯级库区土壤微生物群落特征及其影响因素研究[J]. 长江流域资源与环境, 2023, 32(3): 507-515. (Sun Gan, Sun Ran-hao, Xu Huo-qing, et al. Characterization of Soil Microbial Communities and Their Influencing Factors in Hot-dry Valleys with Cascade Reservoirs[J]. Resources and Environment in the Yangtze Basin, 2023, 32(3): 507-515. (in Chinese))
- [16] 谢益芹, 邓安军, 董先勇, 等. 金沙江下游流域水沙格局变化研究[J]. 水利学报, 2023, 54(11): 1309-1322. (Xie Yi-qin, Deng An-jun, Dong Xian-yong, et al. Study on the Spatial-temporal Variations of Runoff and Sediment in the Lower Reach of Jinsha River[J]. Journal of Hydraulic Engineering, 2023, 54(11): 1309-1322. (in Chinese))
- [17] 龙启建, 李克锋, 汪青辽. 梯级电站联合调节非恒定流对枢纽间航运安全的影响[J]. 水利水运工程学报, 2011(3): 92-97. (Long Qi-jian, Li Ke-feng, Wang Qing-liao. Effect of Unsteady Flow Caused by Joint Regulation of Cascade Hydropower Stations on Safety Navigation between Dams[J]. Hydro-Science and Engineering, 2011(3): 92-97. (in Chinese))

(编辑: 黄 玲)

(上接第 71 页)

- [19] 单义明, 杨 侃. 基于灰色关联度分析的山西省 PSO-SVR 需水量预测模型[J]. 水电能源科学, 2021, 39(2): 18-21. (Shan Yi-ming, Yang Kan. Forecasting Model of PSO-SVR Water Requirement in Shanxi Province Based on Grey Correlation Analysis[J]. Water Resources and Power, 2021, 39(2): 18-21. (in Chinese))
- [20] 王艳菊, 王 珏, 吴泽宁, 等. 基于灰色关联分析的支持向量机需水预测研究[J]. 节水灌溉, 2010(10): 49-52. (Wang Yan-ju, Wang Jue, Wu Ze-ning, et al. Research on Water Demand Prediction of Support Vector Machine Based on Grey Relational Analysis[J]. Water Saving Irrigation, 2010(10): 49-52. (in Chinese))
- [21] 何淑林, 刘慧敏, 金立强, 等. 基于神经网络算法的果树需水预测研究[J]. 灌溉排水学报, 2022, 41(1): 19-24. (He Shu-lin, Liu Hui-min, Jin Li-qiang, et al. Calculating Demands of Fruit Trees for Water Using Neural Network Algorithm[J]. Journal of Irrigation and Drainage, 2022, 41(1): 19-24. (in Chinese))
- [22] 李 雨, 王 君, 张萌萌, 等. 基于 PSO-BP 模型的省域交通运输碳排放多情景预测[J]. 华南师范大学学报(自然科学版), 2025, 57(2): 12-22. (Li Yu, Wang Jun, Zhang Meng-meng, et al. Multi-scenario Prediction of Provincial Transportation Carbon Emissions Using the PSO-BP Model[J]. Journal of South China Normal University (Natural Science Edition), 2025, 57(2): 12-22. (in Chinese))
- [23] 张晓琴, 牛建永, 李顺勇. 基于 G-Q 的 K-S 异方差检验方法[J]. 山西大学学报(自然科学版), 2019, 42(1): 95-104. (Zhang Xiao-qin, Niu Jian-yong, Li Shun-yong. Heteroscedasticity Test Method of K-S Based on G-Q Test[J]. Journal of Shanxi University (Natural Science Edition), 2019, 42(1): 95-104. (in Chinese))

(编辑: 王 慰)