

刘晓伟, 哈明虎, 雷晓辉, 等. 基于改进 GRA-NARX 模型的不同预见期泵站水位预测[J]. 南水北调与水利科技(中英文), 2023, 21(6): 1134-1144. LIU X W, HA M H, LEI X H, et al. Water level prediction for pumping stations with different forecast periods based on improved GRA-NARX model[J]. South-to-North Water Transfers and Water Science & Technology, 2023, 21(6): 1134-1144. (in Chinese)

基于改进 GRA-NARX 模型的不同预见期 泵站水位预测

刘晓伟^{1,2}, 哈明虎¹, 雷晓辉^{1,3}, 张召³, 王超⁴

(1. 河北工程大学水利水电学院, 河北 邯郸 056038; 2. 河北省智慧水利重点实验室, 河北 邯郸 056038;
3. 中国水利水电科学研究院, 北京 100038; 4. 河北工程大学信息与电气工程学院, 河北 邯郸 056038)

摘要: 基于超参数自动率定的 GRA-NARX (grey relation analysis-nonlinear autoregressive model with exogenous inputs) 模型是 GRA-NARX 模型的一种有效改进。以南水北调东线一期工程洪泽泵站为例, 使用基于超参数自动率定的 GRA-NARX 模型, 针对 1 h 和 2 h 时间间隔的输入数据, 分别预测 3 个短预见期(2 h、4 h、6 h)和 1 个长预见期(12 h)的泵站前水位, 并与 GRA-BP (grey relation analysis-back propagation) 模型的预测结果进行比较。结果表明: 不同预见期(2 h、4 h、6 h、12 h)下, 基于超参数自动率定的 GRA-NARX 模型的相关系数、均方根误差和平均绝对误差等指标均相差不大, 预测精度高, 且皆优于 GRA-BP 模型; 采用 1 h 时间间隔的输入数据预测结果优于 2 h 时间间隔的输入数据结果。研究成果可为不同预见期泵站前水位预测提供理论参考。

关键词: 超参数自动率定; GRA-NARX 神经网络; GRA-BP 神经网络; 水位预测; 不同预见期

中图分类号: TV68 **文献标志码:** A **DOI:** 10.13476/j.cnki.nsbdqk.2023.0111

明渠调水工程输水距离长, 泵站、闸门、倒虹吸、分水口等建筑物繁多, 水位受调控建筑物影响较大, 特别是梯级泵站调水工程。当泵站机组切换时, 泵站抽水流量突然发生变化, 导致泵站前水位也随之发生急剧变化, 进而增加了水位预测的难度。为了较好地解决此类问题, 越来越多的智能技术应运而生, 神经网络是其中公认的一种, 已被广泛用于水位预测中^[1-6]。BP 神经网络可以灵活地处理非线性问题并达到较高的预测精度, 但存在局部极小化等问题^[7-8]。长短时记忆(long short term memory, LSTM)网络, 由一个输出门、一个输入门、一个遗忘门和一个记忆单元组成, 它对波动的时间序列有效, 但对长序列依旧存在问题^[9]。Elman 神经网络(Elman neural network, ENN)对任何动态的输入-输出关系都有很高的学习能力。然而, 它用不太值得信赖的学习替代了精简的导数计算^[10]。NARX 神经网络是循环神经网络中的一类, 具有反馈记忆功能,

预测精度高, 被广泛用于复杂的多输入和多输出系统^[11]。NARX 神经网络在如何快速达到神经元之间的连接和输入参数的权重方面, 明显优于其他人工神经网络^[12], 并减少了建立模型的参数数量^[13]。以往关于 NARX 模型的研究多集中在风暴预测、地下水水位预测和旱涝预测等方面。例如: Di Nunno 等^[14]为预测极端风暴事件建立了两个不同的基于 NARX 的模型, 比统计学模型更准确; Di Nunno 等^[15]发现 NARX 神经网络可以精确预测位于意大利 Apulia 的 76 口井的地下水位的波动; Wunsch 等^[16]指出 NARX 神经网络在预测德国西南部断裂、多孔和岩溶含水层地下水水位时很稳定; Ezzeldin 等^[17]揭示了 NARX 神经网络在预测侧孔口的流量方面优于文中的其他模型; Wang 等^[18]证明 NARX 神经网络可以成功地预测中国长江流域的干旱和洪水; Yang 等^[19]应用 NARX 神经网络模拟了位于泰国湄南河流域上游的 3 个多用途水库的运行, 模拟效果好。然而,

收稿日期: 2023-05-05 修回日期: 2023-09-29 网络出版时间: 2023-11-15

网络出版地址: <https://link.cnki.net/urlid/13.1430.TV.20231114.1344.006>

基金项目: 国家自然科学基金项目(52009026); 河北省自然科学基金面上项目(A2020402013)

作者简介: 刘晓伟(1983—), 女, 河北邯郸人, 讲师, 博士研究生, 主要从事复杂水资源系统预报与调度研究。E-mail: 470596294@qq.com

通信作者: 哈明虎(1963—), 男, 河北肃宁人, 教授, 博士, 主要从事不确定信息处理、复杂水资源系统预报与调度研究。E-mail: mhha@hbu.cn

将 NARX 神经网络应用于泵站站前水位预测的研究并不多见。

此外,输入变量的选取对神经网络模型的预测精度有重要影响,应筛选输入变量来降低预测的复杂性、提高预测的精度。灰色关联分析(grey relational analysis, GRA)是 Deng^[20] 在 1982 年提出的,目的是确定决策过程中各解决方案的相关等级,寻找最佳解决方案。近年来,人们通过将 GRA 与神经网络相结合提出了许多预测模型。杨凌雁^[21] 把 GRA 和 BP 神经网络相结合来预测高速公路造价,预测的精度较高;Liu 等^[22] 利用 GRA-LSTM 模型预测了北京市废水排放量,预测精度高于文中其他模型;刘晓伟等^[23] 建立了 GRA-NARX 神经网络模型,用于泵站站前水位预测,预测结果比 GRA-BP 模型好。

值得注意的是,利用已有的神经网络进行预测时,随预见期增长,预测精度变低^[7];以往对 NARX 神经网络中超参数(时间延迟、隐藏层神经元个数)的选取主要依靠研究者的经验进行确定,具有主观任意性,缺乏成熟的算法支撑^[23-26],预测精度不能达到最佳。Liu 等^[27] 对 GRA-NARX 神经网络模型进行了改进,构建了基于超参数自动率定的 GRA-NARX 神经网络模型,将其应用于密云水库调蓄工程屯佃泵站,预测 2 h 水位,结果比 GRA-NARX 模型的预测精度提升 28.3%,比 GRA-BP 模型的预测精度提升 41.5%。但该模型并未考虑不同短预见期、长预见期及不同输入数据时间间隔的泵站站前水位预测。因此,开展不同预见期泵站站前水位预测研究很有意义。

本文运用超参数自动率定的 GRA-NARX 神经网络的泵站站前水位预测模型,分别分析输入数据的时间间隔为 1 h 和 2 h 时,不同预见期水位预测的精度,将模型应用于南水北调东线一期工程洪泽泵站,分别预测 2、4、6 和 12 h 的水位。其中 GRA-NARX 神经网络使用超参数自动率定法得到不同预见期下超参数(时间延迟和隐藏层神经元个数)的最佳组合,并与 GRA-BP 模型的预测结果进行比较,结果可为不同预见期调度提供理论支撑。

1 研究区概况

南水北调东线一期工程江苏段输水干线总长 404 km,连接洪泽湖、骆马湖和微山湖,见图 1。该段工程地形复杂,输水距离长,泵站、闸门、分水口等建筑物繁多,受外界未知因素影响较大,不同调控建筑物之间相互影响,调度运行情况非常复杂。

本文研究对象为南水北调东线一期工程江苏段第 3 级抽水泵站洪泽泵站,工程开工于 2011 年 1 月,竣工于 2013 年 3 月。洪泽泵站的作用有两个:一是抽引金湖站来水入洪泽湖,与淮阴泵站联合运行向北调水,使入洪泽湖流量规模达到 450 m³/s;二是抽排宝应湖、白马湖地区的涝水。洪泽泵站建筑物主要由厂房、进水建筑物和出水建筑物等组成,单机流量 37.5 m³/s,总设计流量为 150 m³/s,设计净扬程为 6.0 m。



图 1 南水北调东线一期工程江苏段

Fig. 1 Schematic diagram of the Jiangsu section of the Eastern Route of South-to-North Water Transfers Project

2 数据和方法

2.1 数据选择

选取 2022 年 9 月 1 日至 2022 年 11 月 30 日洪泽泵站站前水位数据、洪泽泵站站后水位数据、洪泽泵站流量数据、金湖泵站站后水位数据、金湖泵站流量数据等,对洪泽泵站的站前水位进行分析预测。这段时间内水位波动较大,无明显规律性,预测难度大。波动的主要影响因素有季节性变化(11 月份水位较高)、洪泽泵站调控期(2022 年 10 月 9 日至 2022 年 10 月 16 日)、上一站金湖泵站联合运行等。

2.2 灰色关联分析

考虑水力及水流传播时间等因素的影响,不同预见期的泵站站前水位与上、下游一定区域内的水情要素相关,且这些因素对站前水位的影响程度不

同。为减少预测的复杂性,保证预测的精度,需要筛选主要影响因子,作为预测模型的输入量。GRA^[20]可以确定决策过程中各影响因子的相关等级,寻找最佳影响因子。该方法对样本数量和分布规律没有严格要求,具有计算量少且解释性强等优点,在水利、农业、工业等领域均得到了广泛应用。

2.3 NARX 神经网络

NARX 神经网络是循环神经网络中的一类,由

输入层、隐含层和输出层 3 部分组成,其拓扑结构见图 2。NARX 神经网络加入了时间延迟和反馈机制,输出值作为 NARX 结构的一部分反馈给输入值,增强了对历史数据的记忆能力。它还减少了输入参数的数量,从而使模型更有效。如果满足以下条件之一,NARX 神经网络训练过程停止:达到最大运行步数;达到训练算法参数设置值;误差梯度低于最小值。NARX 模型在 MATLAB[®]2020 中实现。

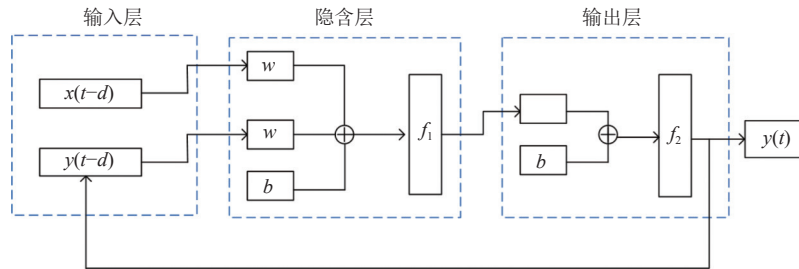


图 2 NARX 神经网络结构
Fig. 2 Structure of NARX neural network

基于 NARX 神经网络的预测模型可表示为

$$y(t) = f[x(t-1), x(t-2), \dots, x(t-d), y(t-1), y(t-2), \dots, y(t-d)] \quad (1)$$

式中: f 为非线性函数; $x(t)$ 表示输入因子; $y(t)$ 表示输出因子; d 表示时间延迟。 $y(t)$ 可由 $x(t)$ 的前 d 个值和 $y(t)$ 的前 d 个值, 通过 f 得到。

2.4 基于超参数自动率定的 GRA-NARX 神经网络

在现有文献^[23-26]中, NARX 神经网络中超参数(时间延迟和隐藏层神经元数量)的选取完全取决于经验。为此, Liu 等^[27]提出了超参数自动率定法,即研究时间延迟和隐藏层神经元个数的所有组合,利用遍历法找出均方误差 E_{MS} 的最小值,确定模型方案, E_{MS} 取最小值时,超参数的组合为最佳组合;构建了基于超参数自动率定的 GRA-NARX 模型,该模型是 GRA-NARX 模型^[23]的有效改进。基于超参数自动率定的 GRA-NARX 模型流程图见图 3。

2.5 预测结果评判

以均方误差 E_{MS} 、平均绝对误差 E_{MA} 、均方根误差 E_{RMS} 和相关系数 R 为评判指标对预测信息精度进行评判。 E_{MS} 反应的是预测值的平方误差, E_{MA} 和 E_{RMS} 反应的是预测值的平均误差, R 反应的是预测值与实际值的线性关系, R 越接近 1, 预测效果越好。

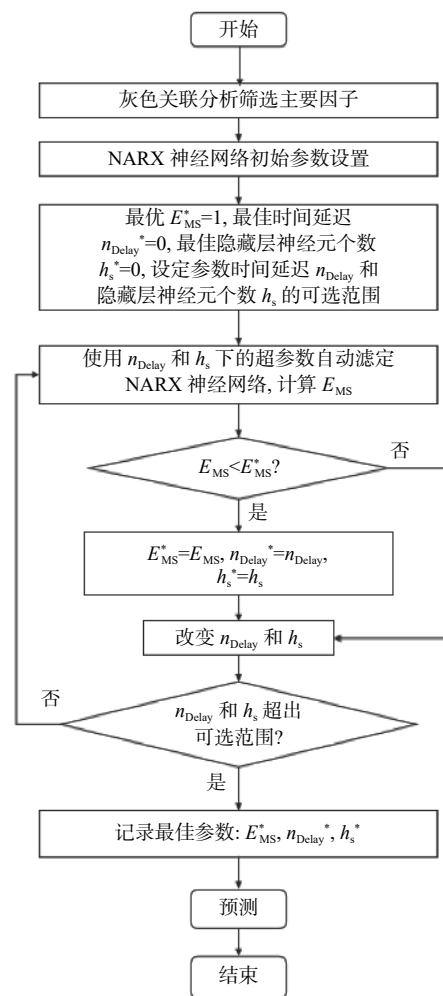


图 3 基于超参数自动率定的 GRA-NARX 神经网络模型流程
Fig. 3 Flow chart of GRA-NARX neural network model of automated hyper-parameter calibration

$$E_{MS} = \frac{\sum_{i=1}^m (f_i - y_i)^2}{m} \quad (2)$$

$$E_{MA} = \frac{\sum_{i=1}^m |f_i - y_i|}{m} \quad (3)$$

$$E_{RMS} = \sqrt{\frac{\sum_{i=1}^m (f_i - y_i)^2}{m}} \quad (4)$$

$$R = \frac{\left| \sum_{i=1}^m (y_i - \bar{y})(f_i - \bar{f}) \right|}{\sqrt{\left(\sum_{i=1}^m (y_i - \bar{y})^2 \right) \left(\sum_{i=1}^m (f_i - \bar{f})^2 \right)}} \quad (5)$$

式中： y_i 和 f_i 分别为泵站站前水位实测值和预测值； \bar{y} 和 \bar{f} 分别为实测水位均值和预测水位均值； m 为数列长度。

3 结果与讨论

3.1 结果

3.1.1 水位影响因子筛选

研究洪泽泵站不同预见期的站前水位时,考虑其可能会受到洪泽泵站当前时刻的流量、当前时刻

站前水位、当前时刻站后水位,当前时刻金湖泵站站后水位以及金湖—洪泽 2 站泵站间流量差等因素的影响。利用灰色关联分析法对 2022 年 9 月 1 日—2022 年 11 月 30 日的 2 162 个监测数据进行分析,结果见表 1。灰色关联度从高到低排序依次为:洪泽泵站当前时刻站前水位、洪泽泵站当前时刻站后水位、洪泽泵站当前时刻的流量、金湖—洪泽 2 站泵站间流量差、当前时刻金湖泵站站后水位。

表 1 各影响因子灰色关联度结果

Tab. 1 Grey relational grade of various influencing factors

序号	因子	关联度
1	洪泽泵站流量	0.799 4
2	洪泽泵站站前水位	0.972 0
3	洪泽泵站站后水位	0.866 1
4	金湖泵站站后水位	0.586 9
5	金湖—洪泽2站泵站间流量差	0.721 7

将灰色关联度大于 0.8 的因子作为神经网络模型的输入因子^[28],以不同预见期的站前水位作为神经网络模型训练阶段的输出因子。

3.1.2 模型参数设置

为对比基于超参数自动率定的 GRA-NARX 神经网络与 GRA-BP 神经网络的性能,所选模型参数见表 2。

表 2 两种模型参数设置

Tab. 2 Parameter settings of the two models

模型	学习率	训练次数	隐含层激活函数	输出层激活函数	训练函数	允许误差	隐含层层数	隐含层神经元个数
基于超参数自动率定的 GRA-NARX模型	10^{-3}	1 000	tansig	purelin	Levenberg-Marquardt	0.000 1	1	超参数自动率定确定
GRA-BP模型	10^{-3}	1 000	tansig	purelin	Levenberg-Marquardt	0.000 1	2	64

3.1.3 水位预测结果分析

分别采用 1 h 和 2 h 时间间隔的输入数据,利用基于超参数自动率定的 GRA-NARX 模型对洪泽泵站站前水位进行 2、4、6 和 12 h 的预测,并与 GRA-BP 模型预测结果进行对比。

1 h 时间间隔输入数据。当输入数据的时间间隔为 1 h 时,依据 2.4 节,设置基于超参数自动率定的 GRA-NARX 神经网络时间延迟的范围为 1~20,隐藏层神经元的个数范围为 1~20,超参数自动率定的 GRA-NARX 模型在不同预见期下超参数(时间延迟和隐藏层神经元个数)不同组合的 E_{MS} 值见图 4。颜色越浅表示 E_{MS} 值越小。图 4(a)显示,当

预见期为 2 h,时间延迟为 19 且隐藏层神经元个数为 12 时, E_{MS} 值达到最小值 0.000 334 1;图 4(b)显示,当预见期为 4 h,时间延迟为 16 且隐藏层神经元个数为 14 时, E_{MS} 值达到最小值 0.000 337;图 4(c)显示,当预见期为 6 h,时间延迟为 17 且隐藏层神经元个数为 9 时, E_{MS} 值达到最小值 0.000 337 8;图 4(d)显示,当预见期为 12 h,时间延迟为 19 且隐藏层神经元个数为 11 时, E_{MS} 值达到最小值 0.000 316 1。由此看出,在不同预见期下, GRA-NARX 模型由于进行了超参数自动率定,最小 E_{MS} 的值相差不大,短预见期与长预见期下的预测精度均很高。

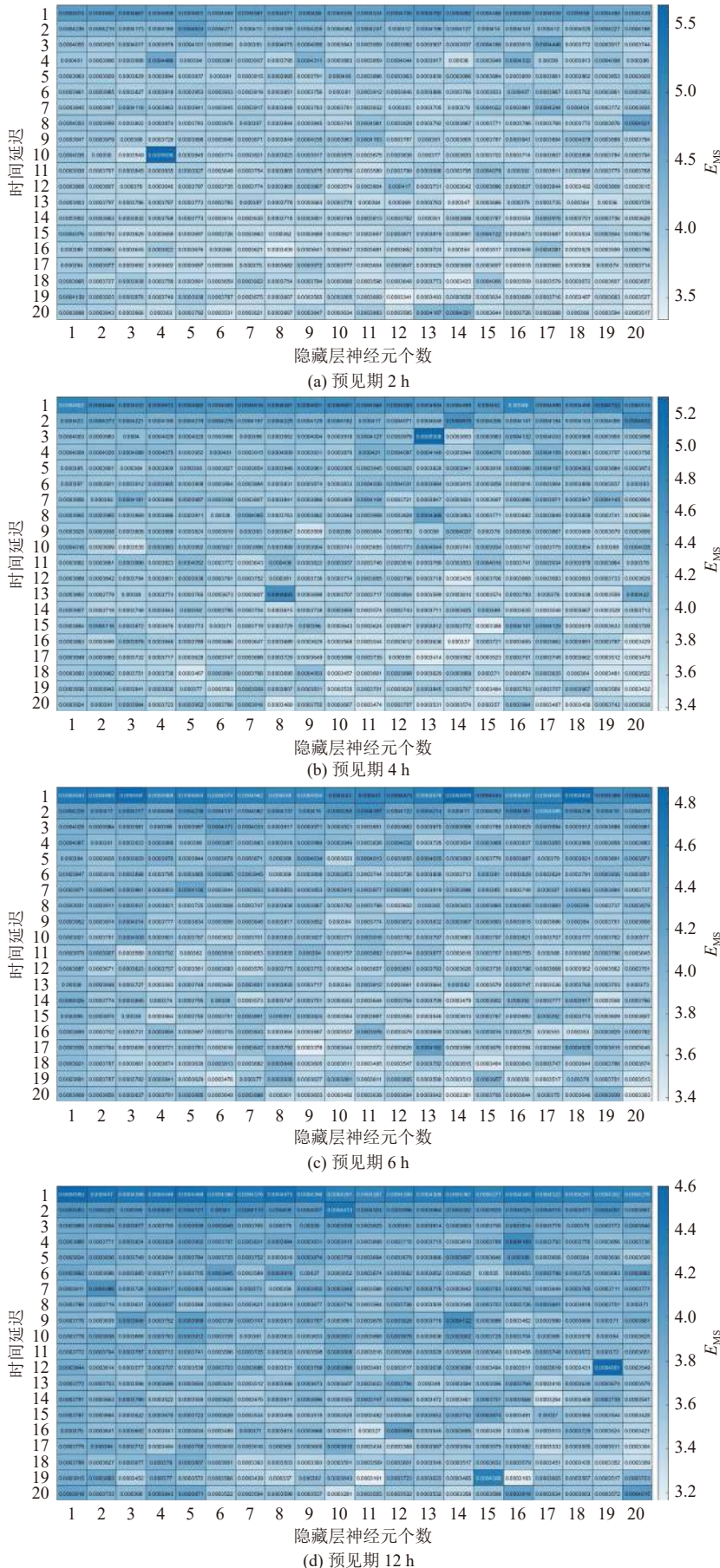


图 4 不同预见期超参数(时间延迟和隐藏层神经元个数)不同组合下均方误差热图

Fig. 4 Heat map of mean square error under different combinations of hyper-parameter (time delay and number of hidden neurons)

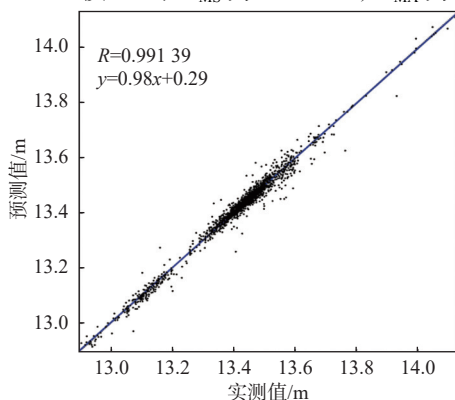
图 5 为超参数自动率定的 GRA-NARX 模型和 GRA-BP 模型在不同预见期下水位实测值与预测值散点图,超参数自动率定的 GRA-NARX 模型的超参数选取最优组合。散点图分布越接近 1:1 对角线,则说明预测值越接近实测数据,可通过相关系数 R 来说明其预测的准确性。当预见期为 2 h 时,超参数自动率定的 GRA-NARX 模型的 R 值为 0.991 39, GRA-BP 模型的 R 值为 0.976 53;当预见期为 4 h 时,超参数自动率定的 GRA-NARX 模型的 R 值为 0.990 79, GRA-BP 模型的 R 值为 0.961 45;当预见期为 6 h 时,超参数自动率定的 GRA-NARX 模型的 R 值为 0.991 16, GRA-BP 模型的 R 值为 0.950 98;当预见期为 12 h 时,超参数自动率定的 GRA-NARX 模型的 R 值为 0.990 03, GRA-BP 模型的 R 值为 0.947 6。由此可知,短预见期与长预见期下,超参数自动率定的 GRA-NARX 模型的 R 值变化不大,均接近 1,预测效果好,且均优于 GRA-BP 模型。GRA-BP 模型随预见期的增长, R 值逐渐变小。

表 3 为超参数自动率定的 GRA-NARX 模型和 GRA-BP 模型在不同预见期下的 E_{MS} 、 E_{MA} 和 E_{RMS} 。结果显示:当预见期为 2 h 时,超参数自动率定的 GRA-NARX 模型的 E_{MS} 为 3.341×10^{-4} , E_{MA} 为 0.0121 m, E_{RMS} 为 0.018 3 m, GRA-BP 模型的 E_{MS} 为 1.1×10^{-3} , E_{MA} 为 0.019 7 m, E_{RMS} 为 0.033 2 m;当预见期为 4 h 时,超参数自动率定的 GRA-NARX 模型的 E_{MS} 为 3.37×10^{-4} , E_{MA} 为 0.012 2 m, E_{RMS} 为 0.018 4 m, GRA-BP 模型的 E_{MS} 为 1.7×10^{-3} , E_{MA} 为 0.025 7 m, E_{RMS} 为 0.041 2 m;当预见期为 6 h 时,超参数自动率定的 GRA-NARX 模型的 E_{MS} 为 3.378×10^{-4} , E_{MA} 为 0.011 9 m, E_{RMS} 为 0.018 4 m, GRA-BP 模型的 E_{MS} 为 2.1×10^{-3} , E_{MA} 为 0.026 9 m, E_{RMS} 为 0.045 8 m;当预见期为 12 h 时,超参数自动率定的 GRA-NARX 模型的 E_{MS} 为 3.161×10^{-4} , E_{MA} 为 0.013 0 m, E_{RMS} 为 0.017 8 m, GRA-BP 模型的 E_{MS} 为 2.3×10^{-3} , E_{MA} 为

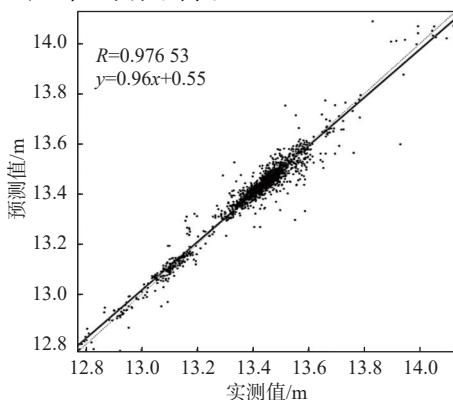
0.030 7 m, E_{RMS} 为 0.048 0 m。通过以上分析说明在短预见期与长预见期下,超参数自动率定的 GRA-NARX 模型的 E_{MS} 、 E_{MA} 和 E_{RMS} 变化不大,比 GRA-BP 模型的 E_{MS} 、 E_{MA} 和 E_{RMS} 都小,预测精度高。GRA-BP 模型的 E_{MS} 、 E_{MA} 和 E_{RMS} 随预见期的增长,逐渐变大,预测精度逐渐变差。

2 h 时间间隔输入数据。当输入数据的时间间隔为 2 h 时,利用超参数自动率定的 GRA-NARX 模型和 GRA-BP 模型分别对洪泽泵站水位进行不同预见期预测,见表 4。结果显示,预见期为 2、4、6、12 h 时,超参数自动率定的 GRA-NARX 模型的 E_{MS} 、 E_{MA} 和 E_{RMS} 值相差不大,分别维持在 0.001 0~0.001 1, 0.019 8~0.021 3, 0.031 6~0.033 2,短预见期与长预见期下均保持较高的预测精度。GRA-BP 模型随预见期的增长, E_{MS} 、 E_{MA} 和 E_{RMS} 值逐渐增大,预测精度变差。不同预见期下,超参数自动率定的 GRA-NARX 模型的预测精度均高于 GRA-BP 模型。与 1 h 时间间隔相比,两模型的预测精度均有下降。

图 6 为超参数自动率定的 GRA-NARX 模型和 GRA-BP 模型在不同预见期下水位实测值与预测值散点图,超参数自动率定的 GRA-NARX 模型的超参数选取最优组合。当预见期为 2 h 时,超参数自动率定的 GRA-NARX 模型的 R 值为 0.969 81, GRA-BP 模型的 R 值为 0.959 35;当预见期为 4 h 时,超参数自动率定的 GRA-NARX 模型的 R 值为 0.967 67, GRA-BP 模型的 R 值为 0.958 25;当预见期为 6 h 时,超参数自动率定的 GRA-NARX 模型的 R 值为 0.970 51, GRA-BP 模型的 R 值为 0.942 79;当预见期为 12 h 时,超参数自动率定的 GRA-NARX 模型的 R 值为 0.969 85, GRA-BP 模型的 R 值为 0.916 67。由此看出,短预见期与长预见期下,超参数自动率定的 GRA-NARX 模型的 R 值变化不大,均优于 GRA-BP 模型, GRA-BP 模型随预见期的增长, R 值逐渐变小。与 1 h 时间间隔比,两模型不同预见期的 R 值均有下降。



(a) 预见期 2 h 的 GRA-NARX 模型散点图



(b) 预见期 2 h 的 GRA-BP 模型散点图

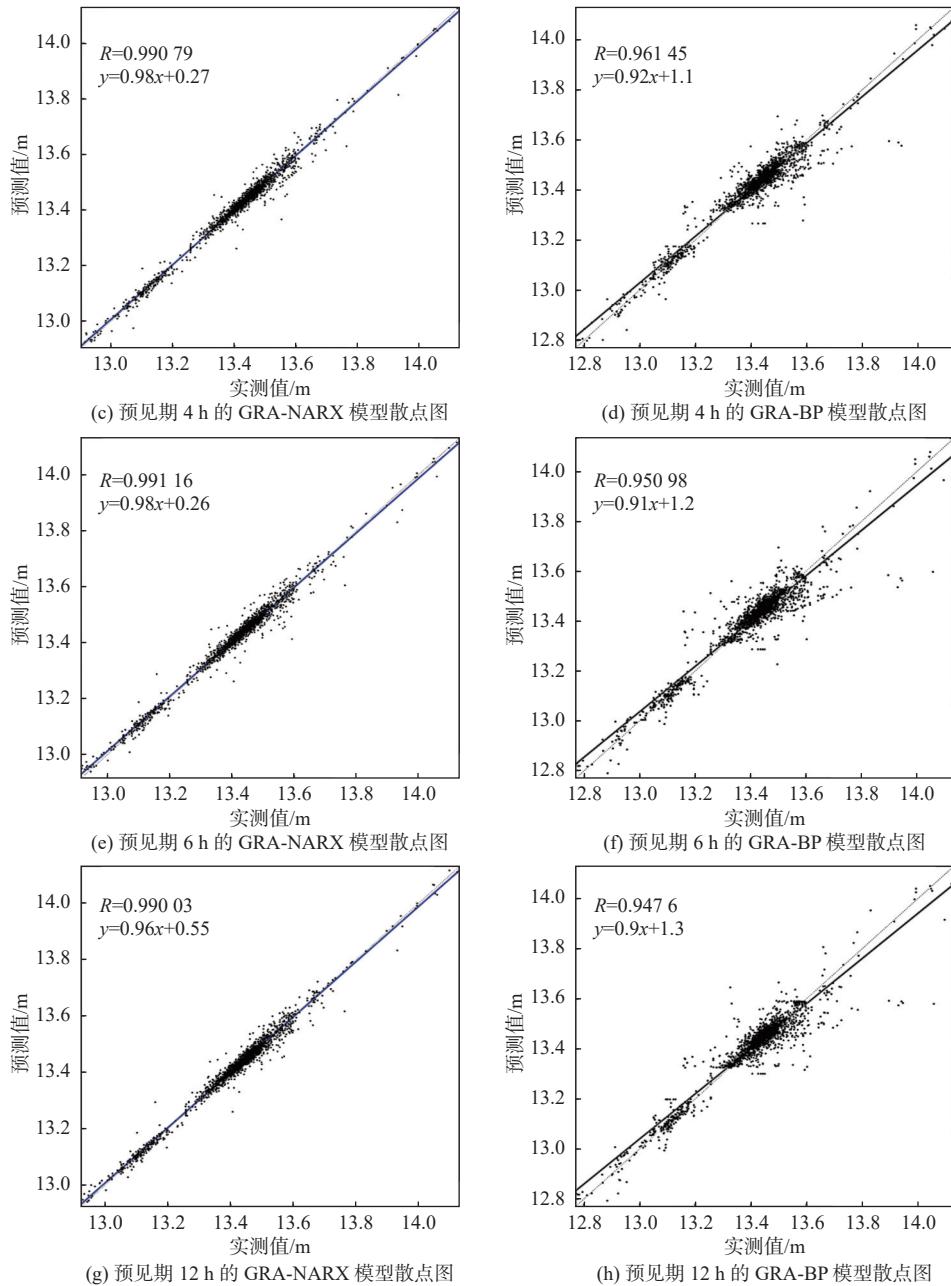


图 5 不同预见期的两种神经网络实测值与预测值散点图(输入数据的时间间隔为 1 h)

Fig. 5 Scatter plots of measured and predicted values for two neural networks of different forecast periods (The data input interval is 1 h)

表 3 两种模型不同预见期的均方误差、平均绝对误差和均方根误差(输入数据的时间间隔为 1 h)

Tab. 3 Mean square errors, mean absolute errors and root mean square errors for two models of different forecast periods (The data input interval is 1 h)

预见期/h	GRA-NARX			GRA-BP		
	E_{MS}	E_{MA}/m	E_{RMS}/m	E_{MS}	E_{MA}/m	E_{RMS}/m
2	3.341×10^{-4}	0.012 1	0.018 3	1.1×10^{-3}	0.019 7	0.033 2
4	3.370×10^{-4}	0.012 2	0.018 4	1.7×10^{-3}	0.025 7	0.041 2
6	3.378×10^{-4}	0.011 9	0.018 4	2.1×10^{-3}	0.026 9	0.045 8
12	3.161×10^{-4}	0.013 0	0.017 8	2.3×10^{-3}	0.030 7	0.048 0

3.2 讨论

研究表明:输入数据的时间间隔为 1 h 或 2 h 时,基于超参数自动率定的 GRA-NARX 模型均可

对短预见期和长预见期的洪泽泵站站前水位进行精准的预测;针对短预见期及长预见期的洪泽泵站站前水位预测问题,基于超参数自动率定的 GRA-

NARX 模型泛化能力要优于 GRA-BP 模型。

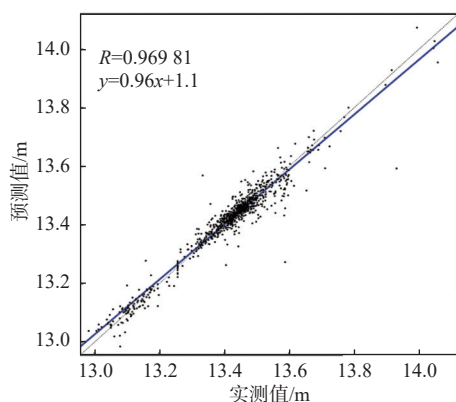
与已有的研究成果^[14-19]相比,本文的研究成果同中有异,相同的是研究结果表明 NARX 神经网络

具备较强的预测能力,不同的是本文利用改进的 NARX 神经网络对短预见期和长预见期泵站水位进行了预测,预测精度均较高。

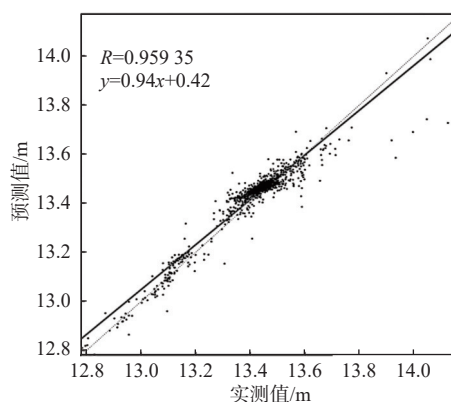
表 4 两种模型不同预见期的均方误差、平均绝对误差和均方根误差(输入数据的时间间隔为 2 h)

Tab. 4 Mean square errors, mean absolute errors and root mean square errors for two models of different forecast periods (The data input interval is 2 h)

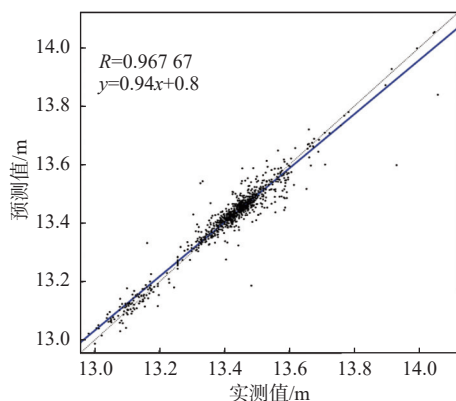
预见期/h	GRA-NARX			GRA-BP		
	E_{MS}	E_{MA}/m	E_{RMS}/m	E_{MS}	E_{MA}/m	E_{RMS}/m
2	0.001 1	0.021 3	0.033 2	0.001 9	0.023 8	0.043 6
4	0.001 1	0.020 5	0.033 2	0.001 9	0.027 0	0.043 6
6	0.001 0	0.019 8	0.031 6	0.002 5	0.029 6	0.050 0
12	0.001 1	0.020 0	0.033 2	0.003 3	0.035 6	0.057 4



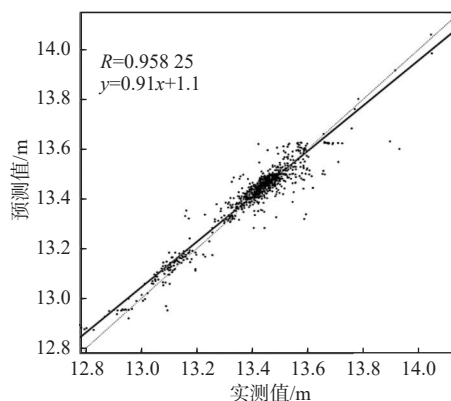
(a) 预见期 2 h 的 GRA-NARX 模型散点图



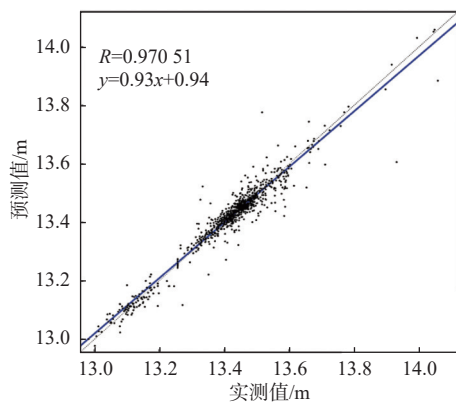
(b) 预见期 2 h 的 GRA-BP 模型散点图



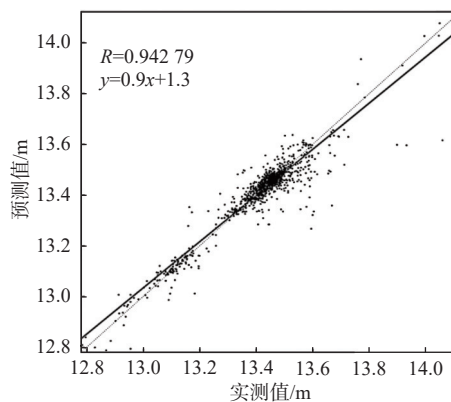
(c) 预见期 4 h 的 GRA-NARX 模型散点图



(d) 预见期 4 h 的 GRA-BP 模型散点图



(e) 预见期 6 h 的 GRA-NARX 模型散点图



(f) 预见期 6 h 的 GRA-BP 模型散点图

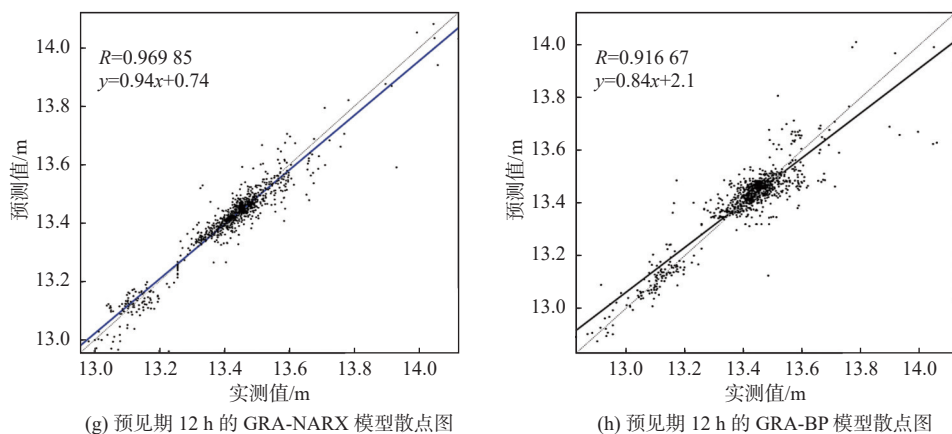


图 6 不同预见期的两种神经网络实测值与预测值散点图(输入数据的时间间隔为 2 h)

Fig. 6 Scatter plots of measured and predicted values for two neural Networks of different forecast periods (The data input interval is 2 h)

今后可进一步考虑模型不同输入因子对预测结果的影响,也可考虑增加分类预测与经典方法的对比,还可与水动力模型进行对比分析。考虑到基于超参数自动率定的 GRA-NARX 神经网络模型对长预见期水位预测精度高,可将其预测结果作为水动力模型的边界条件构建水动力预测模型,为长预见期调度提供理论依据。

4 结论

本文利用基于超参数自动率定的 GRA-NARX 神经网络的泵站站前水位预测模型,分别分析输入数据的时间间隔为 1 h 和 2 h 时,不同预见期水位预测的精度,并与 GRA-BP 模型进行对比。将模型应用于南水北调东线一期工程洪泽泵站,预测 2、4、6 和 12 h 的水位。主要结论如下:

输入数据的时间间隔无论 1 h 还是 2 h,不同预见期下,基于超参数自动率定的 GRA-NARX 模型预测精度均优于 GRA-BP 模型。基于超参数自动率定的 GRA-NARX 模型相关系数、均方误差、平均绝对误差以及均方根误差均相差不大。短预见期下预测精度高,在 12 h 长预见期下仍保持了较高的预测精度。输入数据的时间间隔为 1 h 时,12 h 长预见期下的相关系数为 0.990 03,平均绝对误差为 0.013 m,均方误差为 0.000 316 1,均方根误差为 0.017 8 m。输入数据的时间间隔为 2 h 时,12 h 长预见期下的相关系数为 0.969 85,平均绝对误差为 0.02 m,均方误差为 0.001 1,均方根误差为 0.033 2 m。而 GRA-BP 模型随预见期的增长,相关系数逐渐变小,均方误差、平均绝对误差逐渐增大,预测精度变差。

相同预见期下,输入数据的时间间隔为 1 h 时,超参数自动率定的 GRA-NARX 模型的预测结果优于输入数据的时间间隔为 2 h 时该模型的预测结果,预测精度高。

基于超参数自动率定的 GRA-NARX 神经网络的泵站站前水位预测模型能够满足短预见期和长预见期下泵站水位预测需求,模型还可被应用于其他具有类似结构的调水工程。

参考文献:

- [1] LIU Y, WANG H, FENG W, et al. Short term real-time rolling forecast of urban river water levels based on LSTM: A case study in Fuzhou city China[J]. *International Journal of Environmental Research and Public Health*, 2021, 18(17): 9287. DOI: 10.3390/ijerph18179287.
- [2] 唐鸣, 雷晓辉, 龙岩, 等. 基于长短时记忆网络(LSTM)的南水北调中线水位预测[J]. *中国农村水利水电*, 2020, 10: 189-193. DOI: 10.3969/j.issn.1007-2284.2020.10.034.
- [3] ALSUMAIEI A A. A nonlinear autoregressive modeling approach for forecasting groundwater level fluctuation in urban aquifers[J]. *Water*, 2020, 12: 12030820. DOI: 10.3390/w12030820.
- [4] TU Z J, GAO X G, XU J, et al. A novel method for regional short-term forecasting of water level[J]. *Water*, 2021, 13: 820. DOI: 10.3390/w13060820.
- [5] BARZEGAR R, AALAMI M T, ADAMOWSKI J. Coupling a hybrid CNN-LSTM deep learning model with a boundary corrected maximal overlap discrete wavelet transform for multiscale lake water level forecasting[J]. *Journal of Hydrology*, 2021, 598: 126196. DOI: 10.1016/j.jhydrol.2021.126196.

- [6] REN T, LIU X F, NIU J W, et al. Real-time water level prediction of cascaded channels based on multilayer perception and recurrent neural network[J]. *Journal of Hydrology*, 2020, 585: 124783. DOI: 10.1016/j.jhydrol.2020.124783.
- [7] 薛萍, 张召, 雷晓辉, 等. 基于BP神经网络的不同时间尺度泵站前池水位预测模型[J]. *南水北调与水利科技(中英文)*, 2022, 20(2): 393-407. DOI: 10.13476/j.cnki.nsbdqk.2022.0040.
- [8] 吴美玲, 杨侃, 张铖铖. 基于KG-BP神经网络在秦淮河洪水水位预测中的应用[J]. *水电能源科学*, 2019, 37(2): 74-77,81.
- [9] LIP, ZHANG J, KREBS P. Prediction of flow based on a CNN-LSTM combined deep learning approach[J]. *Water*, 2022, 14: 993. DOI: 10.3390/w14060993.
- [10] LI C, ZHU L, HE Z, et al. Runoff prediction method based on adaptive Elman neural network[J]. *Water*, 2019, 11: 1113. DOI: 10.3390/w11061113.
- [11] CHEN S, BILLINGS S A. Non-linear system identification using neural networks[J]. *International Journal of Control*, 1990, 51: 1191-1214. DOI: 10.1080/00207179008934126.
- [12] DESOUKY M A, ABDELKHALIK O. Wave prediction using wave rider position measurements and NARX network in wave energy conversion[J]. *Applied Ocean Research*, 2019, 82: 10-21. DOI: 10.1016/j.apor.2018.10.016.
- [13] GUZMAN S M, PAZ J O, TAGERT M L M. The use of NARX neural networks to forecast daily groundwater levels[J]. *Water Resources Management*, 2017, 31: 1591-1603. DOI: 10.1007/s11269-017-1598-5.
- [14] DI NUNNO F, GRANATA F, GARGANO R. Forecasting of extreme storm tide events using NARX neural network-based models[J]. *Atmosphere*, 2021, 12: 512. DOI: 10.3390/atmos12040512.
- [15] DI NUNNO F, GRANATA F. GRANATA F. Groundwater level prediction in Apulia region (Southern Italy) using NARX neural network[J]. *Environmental Research*, 2020, 190: 110062. DOI: 10.1016/j.envres.2020.110062.
- [16] WUNSCH A, LIESCH T, BRODA S. Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX)[J]. *Journal of Hydrology*, 2018, 567: 743-758. DOI: 10.1016/j.jhydrol.2018.01.045.
- [17] EZZELDIN R, HATATA L. Application of NARX neural network model for discharge prediction through lateral orifices[J]. *Alexandria Engineering Journal*, 2018, 57(4): 2991-2998. DOI: 10.1016/j.aej.2018.04.001.
- [18] WANG J L, CHEN Y. Using NARX neural network to forecast droughts and floods over Yangtze River basin[J]. *Natural Hazards*, 2022, 110: 225-246. DOI: 10.1007/S11069-021-04944-X.
- [19] YANG S, YANG D, CHEN J, et al. Real-time reservoir operation using recurrent neural networks and inflow forecast from a distributed hydrological model[J]. *Journal of Hydrology*, 2019, 579: 124229. DOI: 10.1016/j.jhydrol.2019.124229.
- [20] DENG J L. Control problems of grey systems[J]. *Systems & Control Letters*, 1982, 1(5): 288-294.
- [21] 杨凌雁. 基于GRA与BP神经网络的工程造价预测模型研究[D]. 武汉: 武汉科技大学, 2015.
- [22] LIU B, WANG S, TANG Y, et al. Prediction of wastewater discharge based on GRA-LSTM: A case study of Beijing[J]. *Environmental Science and Pollution Research*, 2023, 30(9): 24641-24653. DOI: 10.1007/s11356-022-23971-0.
- [23] 刘晓伟, 哈明虎, 雷晓辉, 等. 基于GRA-NARX神经网络的泵站前池水位预测模型[J]. *南水北调与水利科技(中英文)*, 2022, 20(4): 773-781. DOI: 10.13476/j.cnki.nsbdqk.2022.0079.
- [24] 周添一, 徐庆, 刘振鸿, 等. 优化NARX神经网络对时间序列溶解氧的预测[J]. *东华大学学报(自然科学版)*, 2022, 48(2): 105-110,118. DOI: 10.19886/j.cnki.dhdz.2021.0034.
- [25] DI NUNNO F, DE MARINIS G, GARGANO R, et al. Tide prediction in the Venice Lagoon using nonlinear autoregressive exogenous (NARX) neural network[J]. *Water*, 2021, 13: 1173. DOI: 10.3390/w13091173.
- [26] SHAO Y, ZHAO J, XU J, et al. Application of rainfall-runoff simulation based on the NARX dynamic neural network model[J]. *Water*, 2022, 14(13): 2082. DOI: 10.3390/w14132082.
- [27] LIU X W, HA M H, LEI X H, et al. A novel GRA-NARX model for water level prediction of pumping stations[J]. *Water*, 2022, 14: 2954. DOI: 10.3390/w14192954.
- [28] 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, 7: 12,45-48.

Water level prediction for pumping stations with different forecast periods based on improved GRA-NARX model

LIU Xiaowei^{1,2}, HA Minghu¹, LEI Xiaohui^{1,3}, ZHANG Zhao³, WANG Chao⁴

(1. School of Water Conservancy and Hydroelectric Power, Hebei University of Engineering, Handan 056038, China; 2. Hebei Key Laboratory of Intelligent Water Conservancy, Handan 056038, China; 3. China Institute of Water Resources and Hydropower Research, Beijing 100038, China;

4. School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, China)

Abstract: Large-scale water transfer projects make a considerable contribution to reducing a country's uneven distribution of water resources. The water level between two neighboring pumping stations should be kept as consistent as feasible in an open-channel water transfer project that includes a wide range of hydraulic structures for various purposes to prevent potential channel overflow or drying-up of the pumping station forebay. A sharp change in water level may impair water supply and generate significant hydraulic oscillation. As a result, precise prediction of water levels in front of pumping stations is critical to the normal operation of these pumping stations.

The GRA-NARX (grey relationship analysis-nonlinear autoregressive model with exogenous inputs) model is a recurrent dynamic network composed of input delay and feedback memory nodes, with advantageous properties such as more rapid calculation, high generalization performance and high accuracy. The GRA-NARX model based on hyper-parameter automatic calibration is an effective improvement of the GRA-NARX model, which significantly improves the accuracy of water level prediction in front of the Tundian pumping station of the Miyun project in the 2 h short forecast period. However, this model did not consider the water level prediction for different short forecast periods (4 h, 6 h) and long forecast period (12 h). Taking the Hongze pumping station of the Eastern Route of South-to-North Water Transfers Project as an example, the model is used to predict the water level in front of the pumping station for three short forecast periods (2 h, 4 h, 6 h) and one long forecast period (12 h) based on input data at 1 hour and 2 hours intervals, and the prediction results are compared with the GRA-BP (grey relationship analysis backpropagation) model.

The results show that the prediction accuracy of the GRA-NARX model with automatic calibration of hyper-parameters is better than that of the GRA-BP model under different forecast periods. The coefficient of association (R), root mean square error (E_{RMS}), and mean absolute error (E_{MA}) of the GRA-NARX model with automatic calibration of hyper-parameters are not significantly different under different forecast periods. When the time interval for inputting data is 1 hour, with a long forecast period of 12 hours, the optimal R is 0.990 03, E_{RMS} is 0.018 m, and E_{MA} is 0.013 m. When the time interval for inputting data is 2 hours, with a forecast period of 12 hours, the optimal R is 0.969 85, E_{RMS} is 0.033 m, and E_{MA} is 0.02 m.

Since the hyper-parameters of the GRA-NARX neural network are automatically calibrated and the optimal combination of time delay and the number of hidden layer neurons is selected, the GRA-NARX neural network based on the automatic calibration of hyper-parameters under different forecast periods has a good prediction effect on the water level prediction in front of the pumping station, which can meet the water level prediction demand of the pumping station in different forecast periods. Under the same forecast period, when the time interval of inputting data is 1 hour, the prediction results of the GRA-NARX model with hyper-parameter automatic calibration are better than those of the model with inputting data time interval of 2 hours, and the prediction accuracy is high. The findings of the study can be used as the groundwork for estimating the water level in front of pumping stations across different forecast periods.

Key words: hyper-parameter automatic calibration; GRA-NARX neural network; GRA-BP neural network; water level prediction; different forecast period