

# 基于多因子XLSTM模型的地下水位动态模拟与预测

## ——以中国邢台为例

刘 标<sup>1\*</sup>, 张海波<sup>1</sup>, 王 锴<sup>2</sup>, 张心月<sup>3</sup>, 王浩冉<sup>3</sup>, 范 博<sup>3</sup>

<sup>1</sup>新疆邢美矿业有限公司, 新疆 库尔勒

<sup>2</sup>中国煤炭地质总局水文地质局, 河北 保定

<sup>3</sup>中国煤炭地质总局第三水文地质队, 河北 邯郸

收稿日期: 2026年3月4日; 录用日期: 2026年4月21日; 发布日期: 2026年4月29日

### 摘 要

地下水是重要的战略性淡水资源, 在开采过程中出现的大规模地下水位的下降是一个普遍存在的问题, 地下水有效预报对地下水资源管理有着重要意义。本文采用最新提出的xLSTM模型来解决上述难点。本文以河北省邢台地区选取的5口观测孔2018~2023年实际资料为依据, 并根据降水量、温度、蒸发量等因素对模型进行率定分析。上述试验表明, 在模型精度及适应性方面xLSTM模型优于传统的LSTM及BP, 对于水位波动较大的情况具有更好的识别效果, 可以更好地指导可持续开采困难地区的实践工作。

### 关键词

XLSTM, 地下水位, 深度学习, 时序建模

# Dynamic Simulation and Prediction of Groundwater Level Based on Multi-Factor XLSTM Model

## —A Case Study of Xingtai, China

Biao Liu<sup>1\*</sup>, Haibo Zang<sup>1</sup>, Kai Wang<sup>2</sup>, Xinyue Zhang<sup>3</sup>, Haoran Wang<sup>3</sup>, Bo Fan<sup>3</sup>

<sup>1</sup>Xinjiang Xingmei Mining Co., Ltd., Korlai Xinjiang

<sup>2</sup>China Coal Geology Bureau Hydrogeology Bureau, Baoding Hebei

<sup>3</sup>China Coal Geology Bureau Third Hydrogeological Team, Handan Hebei

Received: March 4, 2026; accepted: April 21, 2026; published: April 29, 2026

\*通讯作者。

文章引用: 刘标, 张海波, 王锴, 张心月, 王浩冉, 范博. 基于多因子 XLSTM 模型的地下水位动态模拟与预测[J]. 地球科学前沿, 2026, 16(4): 623-635. DOI: 10.12677/ag.2026.164057

## Abstract

Groundwater is a key freshwater resource, yet long-term extraction has led to widespread declines in water levels, posing challenges for sustainable management. Accurate forecasting is therefore essential. This study applies a recently developed xLSTM model to simulate groundwater dynamics using data from five observation wells in Xingtai, Hebei Province (2018~2023), incorporating precipitation, temperature, and evaporation as driving factors. Results show that, for most wells and metrics, xLSTM generally achieves higher accuracy and more stable performance than LSTM and BP models. It is particularly effective in capturing periods of pronounced water level fluctuations, indicating its potential for supporting groundwater management in complex hydrological settings.

## Keywords

xLSTM, Groundwater Level, Deep Learning, Time Series Modeling

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

地下水是人类生活用水、灌溉用水以及工业生产用水的主要来源, 占全世界灌溉用水的很大比例[1]-[3]; 大量的地下水资源被过量开采也会进一步加剧缺水现状, 使得日常用水得不到保证[4] [5]。近几十年以来, 由于地下水的过度开采, 已经导致了世界范围内的地下水资源的枯竭。地下水水位出现明显的下降趋势[6] [7]。

为了减轻由于抽水引起的现有和潜在的水源短缺, 地下水水位是必须首先监测并且可能预测的一个重要指标[8] [9]。对过去的、现在的以及将来的地下水水位(地下水水位)变化有清晰的认识能够帮助决策者为地区地下水资源可持续发展做出战略规划。

给出地下水变化过程的量化表达。PBM 需要知道丰富的水文地质参数以及边界信息, 一般适用于复杂地下水系统模拟。DDM 基于统计学、机器学习或人工智能的方法从大量的历史数据中挖掘出地下水流场的变化规律。DDM 更注重数据间的相关性及隐藏的关系, 而非具体的物理过程[10]-[12]。基于物理的地下水模型需要详细的含水层参数、边界条件等等, 这些参数的获取代价非常高[13] [14], 在数据稀缺的地区很难校准。在许多实际情况下, 决策者需要对特定井进行稳健的中短期预测, 而不是完全分布的流场, 这促使了数据驱动模型的使用。

自 21 世纪以来, 数据驱动模型(DDMs)逐渐被引入水文研究领域并得到广泛应用, 如人工神经网络(ANNs)、长短期记忆网络(LSTM)以及逻辑回归等方法[15] [16]。随着机器学习(ML)和深度学习(DL)的持续发展, 支持向量机(SVM)、随机森林(RF)、梯度提升模型以及卷积神经网络(CNN)和递归神经网络(RNN)等多类算法相继被用于地下水水位(GWL)预测研究。

但是由于地下水系统的非线性和多时尺度特性, 上述模型对地下水流场复杂的变化仍然存在一定的不足[17]。相比于传统的物理模型, 机器学习方法不需要考虑复杂的水文机理过程, 只需找到解释变量与响应变量之间的统计映射关系就可以实现比较有效的实时预报[18]。

近年来基于模型的方法结合机器学习方法对非线性水文时间序列进行建模得到了较好的拟合效果, 尤其是针对水文时间序列中的周期性和多重尺度特性表现出了较强的适用性[19]。但是, 一些研究证明传

统的 LSTM 在多变量长序列预测上还有着局限性, 比如梯度消失、遗忘等问题。还有不能有效地学习多驱动因素间复杂而长久的依赖关系等诸多不足之处。

为了弥补这一缺陷, 学者们提出了 xLSTM 模型[20], 即扩展了 LSTM 结构, 在原有 LSTM 单元的基础上增加了一个 sLSTM 和 mLSTM, 同时使用指数门控及残差模块来增强其处理长时间序列的能力。综上所述, 在其他领域的长时间序列预测方面, xLSTM 比 LSTM 具有较好的效果, 但在地下水位预测方面的研究较少, 特别是在考虑气候变化和人为开采双重影响条件下 xLSTM 的适用性及优越性尚未得到充分论证。

这项研究的新颖之处不在于提出了一种新的深度学习架构, 而在于将最近开发的 xLSTM 框架应用于地下水位预测——这是一种在以前的水文研究中尚未研究过的应用。具体而言, 这项工作的贡献有三方面: (1) 实施并定制了 xLSTM 架构, 用于在气候和人为影响下进行多元地下水预测; (2) 使用相同的输入变量和评估指标, 对 xLSTM、标准 LSTM、SVG 模型和 BP 模型进行了精度对比; (3) 评估了 xLSTM 在地下水管理背景下的优势、局限性和适用性。

## 2. 研究区域和数据来源

### 2.1. 研究区

研究区位于河北省南部, 西部与山西省相邻, 北部与衡水、石家庄接壤, 东部与山东省相邻, 其界线为卫运河, 南部紧邻邯郸市。地理坐标为东经 113°45'~115°50', 北纬 36°45'~37°48'。

研究区西靠太行山隆起带东与华北平原断裂带接触西部山区构造体系主要为北东向的华夏系断裂带、北北东向的新华夏系断裂带以及相应的北西向的断裂带。研究区约 70%地区为平原区, 平原区面积约为 8766 km<sup>2</sup>。区内平原以滏阳河为分界线, 分别为中部的冲积平原和西部的山前冲积扇平原。山前平原区, 即滏西平原海拔高度一般在 75~40 m 之间, 地面坡度在 1/1000~1/400 之间。该区河流阶地发育, 与中部平原相比, 坡度较大, 河流切割明显。中部冲积平原位于黑龙港地区, 早期因古黄河的长期泛滥, 导致其地形地貌复杂, 沙丘、岗坡、古河道等呈条状分布, 在中间地区形成大量的封闭洼地。

气候类型为温带大陆性季风气候。冬、春两季区内受西伯利亚大陆性气团和蒙古大陆性气团两股气团控制。冬季寒冷干燥, 降水量少; 春季干燥少雨, 偏西风盛行, 风沙严重, 蒸发强度。夏季受太平洋暖流的影响, 降雨量较大, 一年中 80%的降水量集中在夏季, 同时受太行上地形的影响, 降雨分布不均, 山区降雨量大, 东部平原区降雨量相对较小。邢台市多年平均气温为 13.7℃, 冬季气温低, 多年平均值为-0.5℃, 夏季温度高, 多年平均值为 26.3℃。

### 2.2. 数据来源

本文研究区共有 51 处地下水资源监测站点, 利用观测井获得地下水水位数据, 并构成时间序列以及空间序列。地下水水位数据的时间以及空间信息对整个预测结果有着很大的影响, 在一定程度上决定着所用预测模型的形式和算法, 是建立模型的基础。由于受到监测条件的影响, 大部分监测站点的监测周期相对较短, 数据缺失情况严重, 不能很好地进行长序列建模。

**Table 1.** Selected monitoring well coordinates, surface elevation, well depth, and borehole elevation table

**表 1.** 所选监测井地理坐标、地表标高、井深以及孔口标高表

井口	经度	纬度	井深	地面高度(m)	孔深(m)
A (W-521)	114°31'03"	37°00'48"	203.37	67.357	68.357
B (W-529)	115°02'04"	37°23'30"	264.94	24.797	25.127
C (W-532)	115°01'19"	37°04'08"	50	30.919	31.589

续表

D (W-534)	115°35'46"	36°58'18"	282.59	27.981	28.591
E (W-535)	115°29'55"	36°53'24"	289	32.355	32.855

**Table 2.** Impact factor data information table**表 2.** 影响因子数据信息表

数据类型	空间尺度	时间序列	数据源
降水	各观测点处	2018.1~2023.12	<a href="https://doi.org/10.5281/zenodo.3114194">https://doi.org/10.5281/zenodo.3114194</a>
NDVI	各观测点处	2018.1~2023.12	<a href="http://www.geodata.cn">http://www.geodata.cn</a>
蒸散发	各观测点处	2018.1~2023.12	<a href="http://www.geodata.cn">http://www.geodata.cn</a>
气温	各观测点处	2018.1~2023.12	<a href="http://www.geodata.cn">http://www.geodata.cn</a>

结合数据完整性和代表性等因素, 本文选取全部监测点中的 5 处观测资料连续、持续时间长(时间为 2018~2023 年), 并且既有潜水含水层又有承压含水层的监测井作为研究对象, 基本可以代表不同类型水文地质条件下的地下水埋深的变化趋势, 并在此基础上, 利用对应的日尺度水位数据建立地下水平台预报模型。观测井信息如表 1 所示。影响因素包括降水量、蒸散量、消耗量、气温及 NDVI。用水量直接量化了人类活动对地下水资源系统的强度, 并从资源消耗的来源表征了人类活动的强度。因此, 这些点所在县市的用水量被用来表示人类活动的强度。所有因素均从每口井对应位置的 1 km 网格单元中提取, 以确保与监测井的水文响应特征一致。在模型训练之前, 所有变量都已标准化。所选影响系数数据见表 2。

由于很难获得不同井点尺度下的地下水开采量资料, 考虑到区域农业灌溉是引起地下水水位下降的重要因素之一, 在已考虑的气候和生态环境因子之外, 将人类活动的影响用替代指标来反映, 其定义为:

(1) 灌溉季虚拟变量  $I_t$ : 根据区域农业种植制度, 将每年主要灌溉时段(如 4~9 月)定义为 1, 其余月份为 0;

(2) 降水亏缺指数  $D_t$ : 用于表征降水相对不足程度, 定义为  $D_t = P - P_t$ , 其中  $P$  为多年同期平均降水量,  $P_t$  为当月降水量;

(3) 累积降水亏缺指数  $CD_t$ : 用于刻画持续干旱条件下地下水开采压力的累积效应, 定义为  $D_t = \sum_{i=1}^K (P - P_{t-i})$

上述变量用于间接表征农业灌溉与人类用水活动强度, 并与气象因子共同作为模型输入特征参与地下水位预测。

### 3. 模型与方法

#### 3.1. LSTM 模型

LSTM:LSTM 主要由以下四部分构成: 输入门(inputgate), 遗忘门(forgettingate), 输出门(outputgate)和细胞状态(cellstate)。这四个方面共同作用使得 LSTM 可以确定应该在 cell 中保存或者丢失什么信息。在利用上述三个功能后 LSTM 模型可实现每一步对单元进行更新并学会长时的趋势。模型结构示意图见图 1。

以下是 LSTM 单元的标准计算步骤:

$$i_t = \sigma(\hat{i}_t), \hat{i}_t = w_i^\top x_t + r_i h_{t-1} + b_i \quad (1)$$

$$f_t = \sigma(\hat{f}_t), \hat{f}_t = w_f^\top x_t + r_f h_{t-1} + b_f \quad (2)$$

$$o_t = \sigma(\hat{o}_t), \hat{o}_t = w_o^\top x_t + r_o h_{t-1} + b_o \quad (3)$$

$$a_t = \tanh(\hat{a}_t), \hat{a}_t = w_a^\top x_t + r_a h_{t-1} + b_a \quad (4)$$

矩阵  $w_a$ 、 $w_i$ 、 $w_f$ 、 $w_o$  分别对应隐藏状态  $h_{t-1}$  和单元输入、输入门、遗忘门和输出门之间的重复权重。； $r_a$ 、 $r_i$ 、 $r_f$  和  $r_o$  分别表示特征提取过程中  $x_t$  的权重系数、输入门、遗忘门和输出门； $b_a$ 、 $b_i$ 、 $b_f$  和  $b_o$  是相应的偏差项； $f_t$  是遗忘门向量， $i_t$  是输入门向量， $o_t$  是输出门向量， $a_t$  是由输入门调制的输入。 $\tanh$  表示切线双曲函数， $\sigma$  表示激活函数。

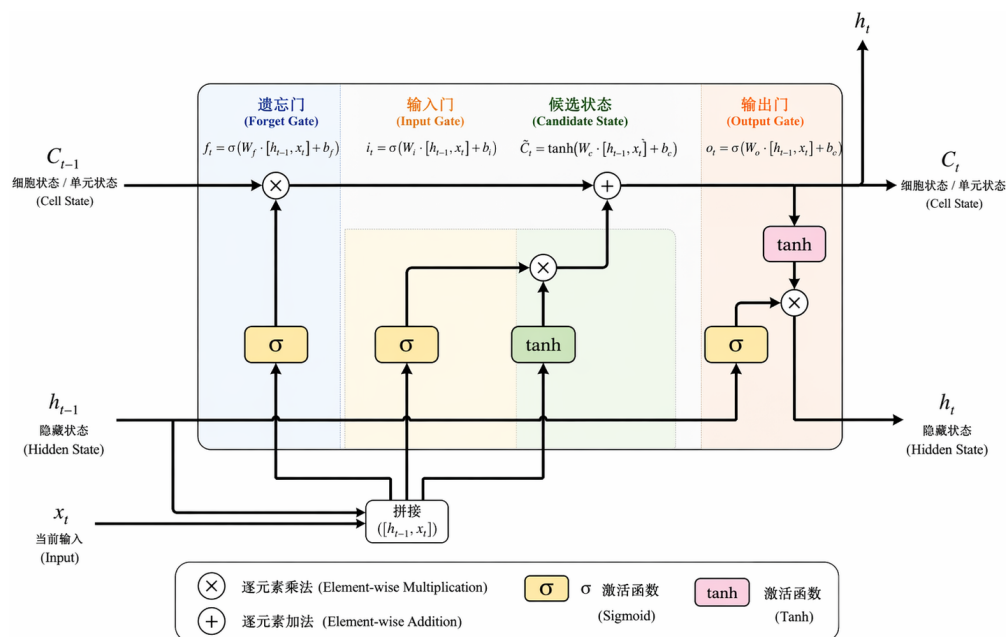


Figure 1. Schematic diagram of LSTM model architecture

图 1. LSTM 模型结构示意图

### 3.2. BP 神经网络模型

BP (Back Propagation)神经网络是 1986 年由 Rumelhart 和 McClelland 提出，是一种按照误差逆向传播算法训练的多层前馈神经网络，是应用最广泛的神经网络。BP 神经网络由输入层、隐含层和输出层构成，其中隐含层有一层或者多层。每一层可以有若干个节点。层与层之间节点的连接状态通过权重来体现。BP 神经网络模型结构见图 2。

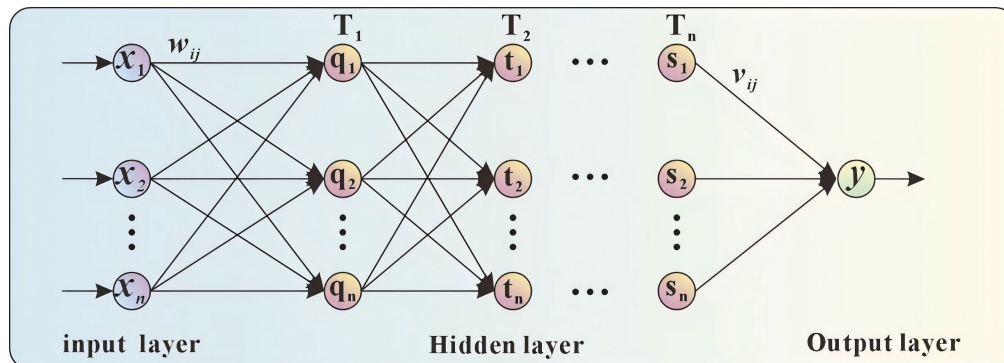


Figure 2. General structure topology diagram of BP neural network model

图 2. BP 神经网络模型通用结构拓扑图

### 3.3. XLSTM 模型

xLSTM 可以看作是 LSTM 的改进及扩展，在面对超长时间跨度的记忆需求时表现出强大的能力。其关键点在于使用了残差单元把 sLSTM 和 mLSTM 结合起来。

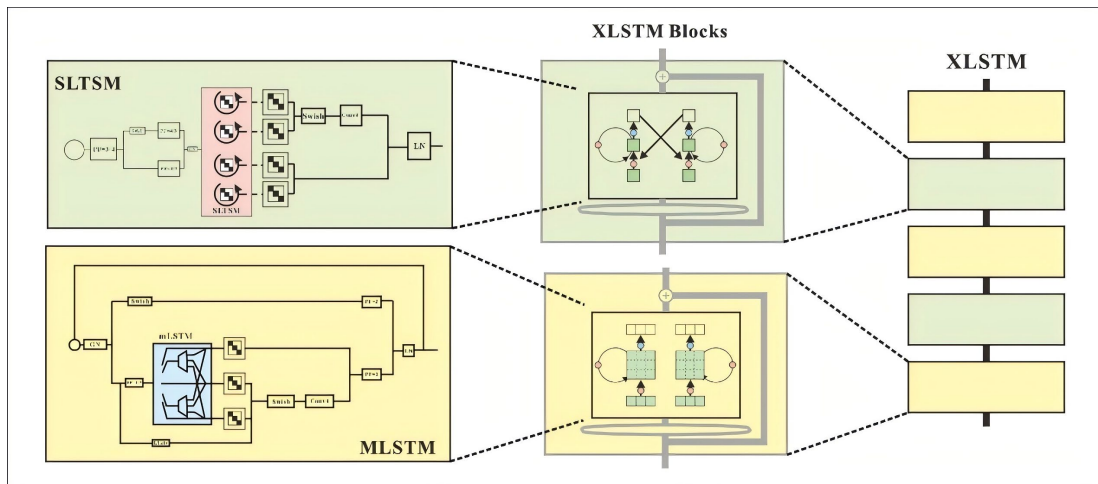
sLSTM：提出一种新的指数级门控(Exponential Gating)以及使用重归一化稳定(Re-Normalized Stability)，通过对输入门及遗忘门的激活函数进行放大来提高网络提取重要信息的能力。sLSTM 的具体实现过程如下所示：

$$z_t = \tanh(\hat{z}_t), \hat{z}_t = w_z^\top x_t + r_z h_{t-1} + b_z \tag{5}$$

$$i_t = \exp(\hat{i}_t), \hat{i}_t = w_i^\top x_t + r_i h_{t-1} + b_i \tag{6}$$

$$f_t = \sigma(\hat{f}_t) \text{ OR } \exp(\hat{f}_t), \hat{f}_t = w_f^\top x_t + r_f h_{t-1} + b_f \tag{7}$$

$$o_t = \sigma(\hat{o}_t), \hat{o}_t = w_o^\top x_t + r_o h_{t-1} + b_o \tag{8}$$



注：1. 引入指数门控的新型 SLSTM 和 mLSTM 存储单元。2. 残留块中的 mLSTM 和 SLSTM 生成 xLSTM 块。3. 堆叠的 xLSTM 块赋予 XLSTM 体系结构。

Figure 3. xLSTM architecture diagram of configuration process

图 3. xLSTM 设置过程体系结构示意图

mLSTM (Matrix LSTM)：摆脱向量化范式的束缚，直接将隐藏状态扩展为一个矩阵，在增强模型记忆能力的同时也增加了运算可并行度；基于层级式同构结构的设计理念，xLSTM 可同时获得各个变体的优势。对于非线性的水文过程而言具有更好的拟合稳定性。mLSTM 的计算方法如下：

$$C_t = f_t C_{t-1} + i_t v_t k_t^\top \tag{9}$$

$$n_t = f_t n_{t-1} + i_t k_t \tag{10}$$

$$h_t = o_t \odot \hat{h}_t, \hat{h}_t = C_t n_t / \max\{|n_t^\top, q_t|, 1\} \tag{11}$$

$$q_t = W_t x_t + b_q \tag{12}$$

$$k_t = \frac{1}{\sqrt{d}} W_k x_k + b_k \tag{13}$$

$$v_t = W_v x_t + b_v \tag{14}$$

$$i_t = \exp(\hat{i}_t), \hat{i}_t = w_i^\top x_t + b_i \quad (15)$$

$$f_t = \sigma(\hat{f}_t) \text{ OR } \exp(\hat{f}_t), \hat{f}_t = w_f^\top x_t + b_f \quad (16)$$

$$o_t = \sigma(\hat{o}_t), \hat{o}_t = w_o^\top x_t + b_o \quad (17)$$

xLSTM 将 sLSTM 和 mLSTM 进行融合并整合在同一个框架中进行设计, 在新的网络架构上有效结合了各改进版本的优势, 提升了模型处理复杂序列信息的能力。xLSTM 设置过程体系结构示意图见 [图 3](#)。

### 3.4. 评估指标

采用 RMSE (Root Mean Square Error)、 $R^2$  (决定系数)与 MAPE (平均绝对百分比误差)来评价模型的整体拟合性能。

Root Mean Square Error (RMSE)是一种用来描述预测值和实际值偏差程度常用的评价指标。其基本思路就是对预测的结果产生多大的误差并且以简单的数值形式表达出来, 这样我们就能知道我们的模型究竟有多精准了, 数值越小, 就越精确; 数值越大, 代表模型误差大。

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^n (Y_i - f(x_i))^2} \quad (18)$$

其中:  $y_i$  为真实值;  $\hat{y}_i$  为模型预测值,  $N$  为样本数。

$R^2$  衡量了模型预测值与真实值之间的关系。其值的范围通常是[0, 1], 表示模型的拟合程度。 $R^2 = 1$ : 完美拟合;  $R^2 = 0$ : 模型没有任何预测能力, 预测值与真实值之间没有线性关系。 $R^2 < 0$ : 模型表现极差, 甚至为欠拟合。

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (19)$$

其中:  $y_i$  为真实值;  $\hat{y}_i$  为模型预测值;  $\bar{y}$  为真实值的均值;  $N$  为样本数。

平均绝对百分比误差(MAPE, Mean Absolute Percentage Error), 是衡量回归模型拟合程度的一个数值指标, 表示的是预测值与真实值之间的平均绝对百分比误差, 它是将每一个预测结果的绝对百分比误差求出后进行平均得到的结果。结果用%表示, MAPE 越小表明所建模型具有较高的拟合度; MAPE 越大则表明模型偏差较大, 即模型预测效果不好。

计算式如下:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (20)$$

其中:  $y_i$  为第  $i$  个真实值;  $\hat{y}_i$  为第  $i$  个预测值;  $N$  为样本数。

### 3.5. 预测不确定性量化方法

采用 RMSE (Root Mean Square Error)、 $R^2$  (决定系数)与 MAPE (平均绝对百分比误差)来评价模型的整体拟合性能。

为了提高模型结果应用于地下水资源风险评估及管理决策的适用性, 在获得点预测的结果基础上再建立预测区间来表征模型的不确定性是必要的。由于本研究中所选取的样本数量较少, 并且各监测井地下水动态变化具有较大的异质性, 因此采用基于前向链式验证(forward chaining)残差分布经验预测区间的不确定性分析方法。在每一轮前向链式验证中记录模型预测残差  $e_t = y_t - \hat{y}_t$ , 并在各监测井测试阶段统计

残差的经验分布, 分别提取其 2.5%与 97.5%分位数, 构建 95%预测区间:  $PI_t = Y_t + q_{0.025}, Y_t + q_{0.975}$ 。

其中,  $Y_t$ 为时刻  $t$  的预测值,  $q_{0.025}$ 和  $q_{0.975}$ 分别为残差经验分布的下、上分位数。为评价区间预测效果, 进一步引入区间覆盖率(PICP)与平均区间宽度(MPIW)两个指标。PICP 反映观测值落入预测区间的比例, MPIW 表征预测区间的平均宽度。PICP 越高且 MPIW 越适中, 说明模型在保证可靠性的同时具有较好的区间判别能力。该方法能够在不显著增加模型复杂度的前提下, 为地下水位预测提供可解释的不确定性范围。

## 4. 结果与讨论

文中使用的模型均为 PyTorch 深度学习框架下的模型。由于每个地下水井的水文地质条件不完全相同, 为了消除空间异质性对模型训练的影响, 文中对每一个监测单独进行建模。模型训练之前, 将所有的输入变量进行标准化, 使其线性变换到 0, 1 区间, 加快模型的收敛过程, 并增加模型训练稳定度。

为了能够公平地比较 xLSTM 和 LSTM, 我们将它们设置成相同的网络结构及重要的超参数。使用 Adam 进行训练, 它结合了动量法和 RMSProp 的优点, 并根据梯度的一阶矩和二阶矩估计来自动调整学习率, 以加快模型的训练速度及保证训练过程的稳定。另外, 选用 MSE (Mean Square Error)损失函数来衡量预测值与实际值之间的误差大小。

### 4.1. xLSTM 模型性能与 LSTM 模型性能

例如 W-535 观测站, BP 神经网络有 3 个隐藏层,  $batch\_size = 32$ , 在 lstm 和 xlstm 模型中, 将隐藏层数设置为 2 来避免过度适应。使用网格搜索方法对超参数进行调优, 并根据隐藏层数为 2 的情况确定  $batch\_size$  大小以及隐藏层大小。Batch\_Size 在[1, 3, 6, 9, 12, 15]中选取, 隐藏层大小的数量从集合[22、23、24、25、26、27、32、64、128、256]中选择。以该模型为例, epochs, learning\_rate 分别设置为 150, 0.0234; 当批大小和隐藏大小分别为 2 和 24 时, 该模型表现最佳。在选择模型变量时, 考虑了降雨、温度和蒸发等因素。我们尝试了不同的比率来划分训练集和测试集, 最终选择了 4:1 的比率, 这最适合(W-535)观测站的水位数据。这确保了训练集有足够的影响因素, 也保证了足够的样本。其他模型中的参数和变量也有类似的规则, 这里不再重复。W-521 观测孔水位拟合结果见图 4。

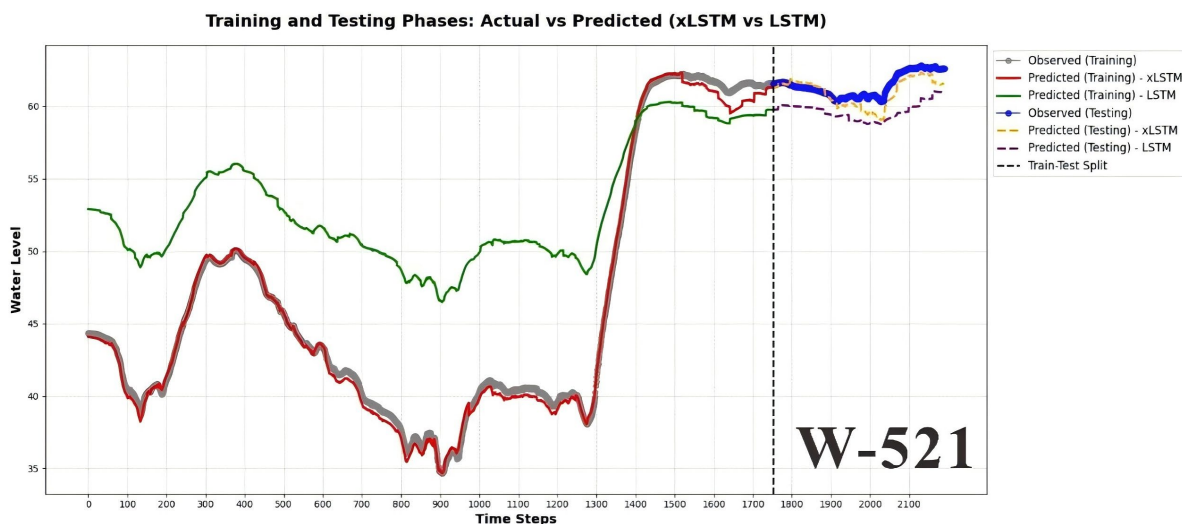


Figure 4. Figure of water level fitting results of observation hole W-521

图 4. W-521 观测孔水位拟合结果图

## 4.2. BP 神经网络模型性能

另建立 BP 神经网络模型对 5 个观测井的地下水水位。模拟试验期为 2018 年至 2023 年，其中训练期与模拟期按照各观测井水位不同需求进行划分。从模型性能方面来看，BP 神经网络模型在捕捉水位变化趋势上稍显不足，特别是在周期性表现不强的观测孔上。以 W-521 为例， $R^2$  在该模型测试期评价指标下降幅度巨大， $R^2$  下降至 0.4136，且从模型效果来看，未能准确捕捉在生态修复时水位变化规律。

在 W-521 井，xLSTM 比 LSTM 和 BP 更接近观测到的突然上升。然而，这仅代表一个异常事件，并不构成对模型在快速水位变化下性能的系统验证。xLSTM 对广义突变预测的适用性需要更严格的评估程序，如移动窗口验证或多事件异常测试。拟合结果对比见图 5。

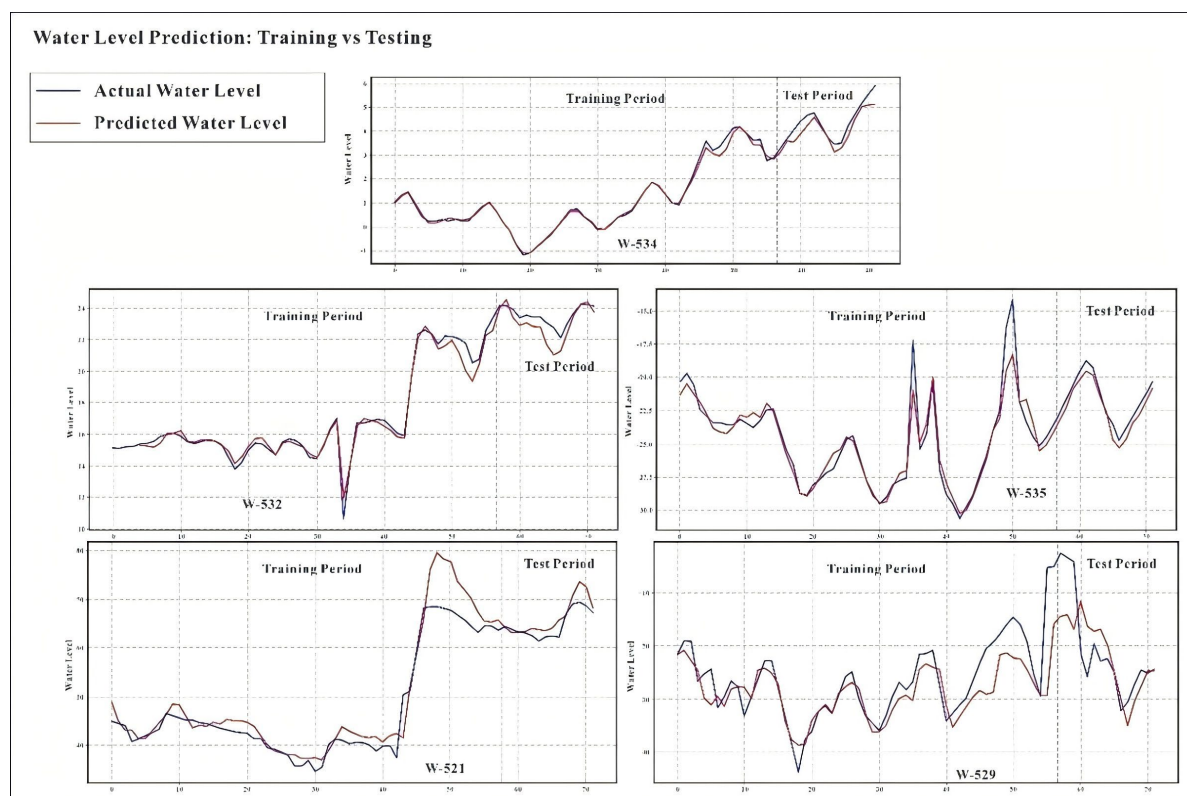


Figure 5. Comparison of LSTM and XLSTM fitting results

图 5. XLSTM 与 LSTM 拟合结果对比图

## 4.3. 模型性能评估

建立的地下水位预测算法性能通过在训练和测试数据集上使用采用 RMSE、 $R^2$  与 MAPE 指标进行了评估。考虑了三种不同的模型：单一 LSTM、XLSTM 和 BP 神经网络。采用 RMSE、和  $R^2$  和 MAPE 指标的预测结果如表 3 所示。

Table 3. Summary of the results of predicting the 5 wells by BP neural network, LSTM and XLSTM

表 3. 使用 BP 神经网络、LSTM 与 XLSTM 预测 5 个井的结果汇总表

Method	Metric	W-521		W-529		W-532		W-534		W-535	
		Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
LSTM	$R^2$	0.9996	0.6046	0.9975	0.9919	0.9980	0.7758	0.9967	0.9382	0.9995	0.9959

续表

XLSTM		0.9999	0.9771	0.9931	0.9829	0.9984	0.9583	0.9988	0.9916	0.9995	0.9984
BP		0.7814	0.4136	0.5581	0.4538	0.9787	0.536	0.9896	0.7517	0.9143	0.8872
LSTM	MAPE	0.0035	0.0045	0.1697	0.2063	0.0041	0.2939	0.0253	0.0017	0.0043	0.0909
XLSTM		0.0017	0.0015	0.1557	0.3729	0.0035	0.0919	0.0131	0.0018	0.0027	0.0334
BP		0.0486	0.024	0.2301	0.8642	0.0214	0.0149	0.104	0.0716	0.0294	0.0387
LSTM	RMSE	0.1823	0.1226	3.7341	5.6516	0.1252	0.3026	0.0814	0.1842	0.0585	0.1165
XLSTM		0.0991	1.8774	3.7247	5.246	0.1149	0.1304	0.0484	0.0679	0.0591	0.0729
BP		3.629	2.1065	4.8771	6.8277	0.73	0.4726	0.1196	0.3736	1.0003	0.5856

结果表明,使用 XLSTM 构建的模型拟合效果优于 LSTM 构建的模型,各观测井  $R^2$  与 MAPE 值均为 XLSTM 模型更优,以 W-535 观测井为例,xLSTM 模型在训练集上的  $R^2$  值为 0.9668,相较于 LSTM 的 0.9209 和 BP 神经网络的 0.9143,分别提高了 5.0%和 5.8%;在测试集上的  $R^2$  值为 0.9034,较 LSTM (0.8477)提升 6.6%,较 BP 神经网络(0.8872)提升 1.8%。

表明 xLSTM 在拟合度以及追踪趋势的能力更强,在预测误差上,xLSTM 模型 MAPE 值在训练集和测试集上的值分别为 0.0176 和 0.0334,远小于 LSTM (0.0271、0.0909)和 BP 神经网络(0.0294、0.0387),其对应的误差降低程度分别是:训练集 35.0%和 39.8%,测试集 63.3%和 13.7%。

这说明 xLSTM 模型不管是短期拟合还是长期预测都比其它模型准确度高;虽然 xLSTM 在测试集 RMSE 上是 0.8256 大于 BP 神经网络(0.7265)和 LSTM (0.8119),但是  $R^2$  和 MAPE 这两个评价指标上却优于 BP 神经网络和 LSTM;BP 神经网络虽然在 RMSE 上略胜于 xLSTM,但是  $R^2$  和 MAPE 均远远低于 xLSTM 的结果,表明其总的拟合效果及预测性能不稳定。整体性能对比分析见表 4。

**Table 4.** Overall performance comparison table  
**表 4.** 整体性能对比分析表

模型	平均训练 $R^2$	平均测试 $R^2$	平均训练 MAPE	平均测试 MAPE	平均训练 RMSE	平均测试 RMSE	稳定性 (训练 vs 测试)
LSTM	0.9209	0.6737	0.0271	0.0909	1.2638	1.9246	差异较大
XLSTM	0.9668	0.7586	0.0176	0.0334	1.2696	1.7351	较好
BP	0.9143	0.6085	0.0294	0.0387	1.7539	2.0731	不稳定

从结果来看,LSTM 模型与 BP 神经网络在预测地下水位方面达到了合理的准确性。相比之下,XLSTM 模型显示出增强的预测准确性。XLSTM 模型优于单 LSTM 模型与 BP 神经网络模型,展示了该算法在提高地下水位预测方面的有效性。从整体性能上来看,XLSTM 模型稳定性大多都优于其余模型,且模拟结果较优,每个站点最优模型见表 5。

**Table 5.** Optimal model and reason table of each site  
**表 5.** 各站点最优模型及原因表

站点	更优模型	理由
W-521	XLSTM	测试 $R^2$ (0.7396)最优
W-529	XLSTM	相较之下, XLSTM 表现较优
W-532	BP	模型构造简单,且表现优异

续表

W-534	XLSTM	比 LSTM 的 RMSE 降低 31% (0.42→0.29)
W-535	XLSTM	测试 R <sup>2</sup> 达 0.9034

地下水位监测数据集中的异常值通常是由监测设备错误和非标准监测方法引起的并且经常介于事实和错误之间。事实表示该值是事实，而错误表示监控方法可能不规则。此外也很难处理。本文利用统计数据的箱线图对各测点的水位时程进行分析，由图 6 可知，测点 W-529 存在一些明显的大于均值的正异常值。

对研究区 5 口观测井采用 LSTM、BP 神经网络以及 XLSTM 模型进行模拟预测，计算了测试期间每口井的 RMSE 和 NSE，并进行了比较。

本研究所用 xLSTM 模型表现出来的相对较好结果，与近些年的相关研究成果大体相合，表明引入了改进结构的循环神经网络用于描述非线性的水文动力学过程具有一定的优越性。相对于传统的模型而言，本文在方法上采用了较为复杂的递归结构(xLSTM)，同时融入了较多的气候因子及人文活动因子，由此在一定程度上提高了模型对地下水位变化过程的描述水平。不过需要说明的是，这种模型效果改善并不是全部监测井都相同。

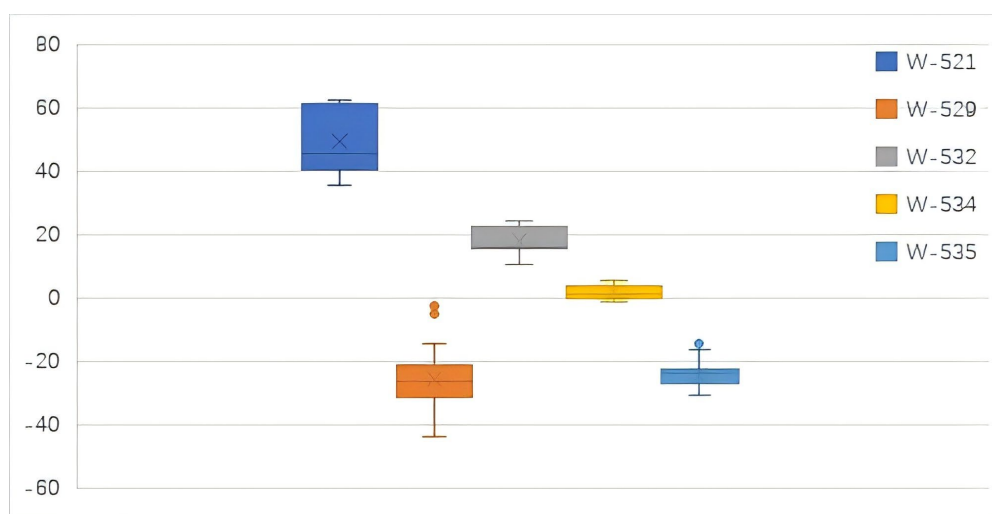


Figure 6. Chart of fluctuation range of water level at each station

图 6. 各站点水位波动范围图

有文献表明，复杂模型的优越性在很大程度上取决于数据自身的丰富程度及数据的变化特征，在输入数据信息量不足或者自身波动特征不够显著的情况下，增加模型结构复杂度所获得的益处是有限的，这也从侧面解释了为何本文中有部分井点预测精度提高不大这一现象

xLSTM 模型大大提高了预测精度，但是仍然有很多未来工作可以提高模型的表现。一种方法是将它与其他模型结合起来。XLSTM 模型存在进一步改进空间，将其与 EMD 结合可以进一步提高模型的预测精度。

这样就可以对短期变化进行预测以及对未来的变化做出规划，从而积极地干预尽可能提高在不同地区的地下水可持续性的可能性。这些方法的效果通常取决于充足且优质的数据；因此，在监测不一致或历史数据有限的地区，其性能可能会下降。本研究中开发的方法论可以调整并应用于这些场景，可能会在水文预测领域带来重大进展。

在 W-521 井, xLSTM 比 LSTM 和 BP 更接近观测到的突然上升。这仅代表一个异常事件, 并不构成对模型在快速水位变化下性能的系统验证。xLSTM 对广义突变预测的适用性需要更严格的评估程序, 如移动窗口验证或多事件异常测试。

#### 4.4. 不同水文情境下模型性能差异及其机制解释

对于丰水期的补给来说, 由于降水入渗补给存在一定的滞后期和积累效应, 在到达含水层之前会经过包气带的运移以及含水层中的传导作用; 而 xLSTM 中增强了门控的作用可以更好地捕捉这种延时的信息, 并且其矩阵的记忆方式也可以较好地模拟气候条件与生态环境因素之间的相互影响, 因此往往在丰水期的表现较佳。故此在补给为主时期优于 LSTM 模型及 BP 模型。

对于枯期或者抽灌耗水期间, 地下水流场的变化呈平缓趋势, 序列变化幅度小, LSTM 和 xLSTM 均能得到较高的精度, 在控制要素较少的情况下 BP 模型也能得到可以接受的结果。可见复杂模型优势并不明显, 仅适用于非线性较强或者多因子耦合程度高的情况。

在水位剧变阶段, 各个模型误差均有所增大, 但是 xLSTM 还是比较有动态响应的能力, 并不逊色于其它模型, 这可能与它的指数门控以及残差连接可以较好地保存突变的信息和深层次的特征有关; 但是这一时期往往由于突发抽采或者异常补给等原因的影响, 现有的变量很难完全刻画, 所以这种优势应该是一个相对优势。

总的来说, xLSTM 更适合于刻画长期依赖、滞后补给及多因子耦合的过程, 但受限于输入信息的完整性和井点的差异性; 在变化平稳的情况下, 简单的模型已经足够高效, 在过程复杂或者驱替因素耦合作用明显的情况下 xLSTM 具有更高的优势。

## 5. 结论

本文基于 2018~2023 年邢台地区 5 口监测井地下水位及多源环境因子数据, 构建了 xLSTM、LSTM 和 BP 神经网络模型, 并采用前向链式验证方法对地下水位动态预测能力进行了对比分析。研究表明, 在本研究样本范围和输入条件下, xLSTM 在多数井点和多数评价指标上表现出较好的预测精度与泛化稳定性, 说明其在表征地下水位序列中的非线性关系、时序依赖及多因子耦合特征方面具有一定优势。

从情景适应度看, 在补给滞后显著、过程阶段性转换频繁、多因子耦合效应明显的时段往往有 xLSTM > LSTM > BP 的趋势, 这是其增效门控、矩阵记忆以及残差连接所致; 但是并不总是如此, 在一些过程平稳、控制机理相对简单的或者局部扰动不易为已有变量所表征的井点, 简单模型还可能取得更好的结果。W-532 井 BP 更优的现象也表明不同的模型对于不同的地下水动态模式有着各自的使用范围。

此外, 在此基础上, 利用前面所得到的前向链式验证残差来构造经验预测区间, 初步刻画了模型中的不确定性, 并发现预测误差随时间变化并不均匀, 出现剧烈波动的时间以及发生驱动变换的时间对应的预测区间较宽, 因此使用点预测可能会低估真实的风险。

同时也要注意, 本文尚存在一些不足之处: 由于所用的监测井个数以及每个监测井对应的监测序列长度都较少, 在一定程度上会影响所得结果的有效性和可靠性; 将人类活动的影响用替代指标来表示, 不能准确地代表井点范围内的抽采行为; 对于未来的预测是依据残差的经验分布给出的, 对可能性大小的描述过于笼统。在今后的工作中可以考虑增加更多观测资料, 并结合更高精度的抽采数据, 对不确定性进行更为全面的概率化预报。

## 参考文献

- [1] 曹文庚, 郭华明, 郑焰, 等. 华北平原地下水砷异常识别-成因-预警关键技术[J/OL]. 地球学报, 1-11.

- <https://link.cnki.net/urlid/11.3474.P.20260326.1426.006>, 2026-04-28.
- [2] 王铂淮, 周帅, 李紫妍, 等. 地下水超采区地下水时空动态演变特征预测及其驱动力研究[J/OL]. 水电能源科学, 1-10. <https://doi.org/10.20040/j.cnki.1000-7709.2026.20251223>, 2026-04-28.
- [3] 李传科, 徐扬, 鲁帆. 基于逐步回归的地下水超采区水位预测及控制指标研究[J]. 中国水利水电科学研究院学报(中英文), 2023, 21(4): 360-368.
- [4] 王军进, 张洪伟, 张国珍, 等. 地下水数值模拟方法的研究与应用进展[J]. 环境与发展, 2018, 30(6): 103-104+106.
- [5] 陈霞, 陈少华, 金震, 等. 地下水污染数值模拟与应用进展[J]. 安徽化工, 2025, 51(4): 120-122.
- [6] 刘玲, 魏亚强, 陈坚, 等. 动网格在非饱和和界面数值模拟中的应用研究进展[J]. 地质科技通报, 2023, 42(1): 360-368+377.
- [7] 王辰辰. 南水北调中线滏阳河段生态补给地下水数值模拟研究[D]: [硕士学位论文]. 邯郸: 河北工程大学, 2025.
- [8] 惠磊, 乔鑫, 高国琰, 等. 基于 MODFLOW 与 CNN-LSTM-Attention 模型的月牙泉水位预测模拟[J/OL]. 水电能源科学, 1-12. <https://link.cnki.net/urlid/42.1231.TK.20260306.0933.006>, 2026-04-28.
- [9] 周宇. 基于深度学习与数值模拟的河谷雨洪资源优化利用研究[D]: [硕士学位论文]. 三河: 防灾科技学院, 2025.
- [10] 饶庆阳, 杨琼波, 崔东文. 基于 WPT 二次分解与 CPO 优化的 KAN 地下水位预测模型[J/OL]. 人民珠江, 1-12. <https://link.cnki.net/urlid/44.1037.TV.20260105.1427.002>, 2026-03-17.
- [11] 潘思成, 潘秀昌, 崔东文. 基于数据处理与超参数优化的高斯过程回归月地下水位预测[J/OL]. 三峡大学学报(自然科学版), 1-8. <https://link.cnki.net/urlid/42.1735.TV.20251111.1704.002>, 2026-03-17.
- [12] Seidu, J., Ewusi, A., Kuma, J.S.Y. and Ziggah, Y.Y. (2026) Optimised Stacking Generalisation Methodology for Groundwater Level Prediction. *Environmental Research Communications*, **8**, Article 035006. <https://doi.org/10.1088/2515-7620/ae49f4>
- [13] Zhang, Y., Li, H., Zhong, Y., Liu, W., Chen, S., Zhang, X., *et al.* (2026) Comparative Assessment of Machine-Learning Models for Daily Groundwater Level Prediction in a Metropolis, Southwestern China. *Journal of Hydrology: Regional Studies*, **64**, Article 103233. <https://doi.org/10.1016/j.ejrh.2026.103233>
- [14] Yashooa, N.K., Boo, K.B.W., Cherubini, C., Huang, Y.F., Sham, F.F., Sherif, M., *et al.* (2026) Integration of an Adaptive Neuro-Fuzzy Inference System (ANFIS) Model for Groundwater Level Prediction Utilizing Feature Engineering Techniques in the Training Process. *Neural Computing and Applications*, **38**, Article No. 79. <https://doi.org/10.1007/s00521-025-11761-x>
- [15] Wei, H., Wei, G., Yu, B., Peng, Y., Wang, M., Xu, B., *et al.* (2026) A Coupled Spatial Reduction-Reconstruction and LSTM Framework (SRR-LSTM) for Groundwater Level Prediction in Large Irrigation Districts. *Scientific Reports*, **16**, Article No. 7450. <https://doi.org/10.1038/s41598-026-37618-4>
- [16] Banadkooki, F.B., Ghanbari-Adivi, E., Sayyahi, F. and Ehteram, M. (2026) An Intelligent Hybrid Deep Learning-Machine Learning Model for Monthly Groundwater Level Prediction. *Scientific Reports*, **16**, Article No. 4132. <https://doi.org/10.1038/s41598-025-34292-w>
- [17] Hu, S., Du, M., Yang, J., Liu, Y., Tuo, Z. and Ma, X. (2025) Application of a Hybrid CNN-LSTM Model for Groundwater Level Forecasting in Arid Regions: A Case Study from the Tailan River Basin. *ISPRS International Journal of Geo-Information*, **15**, Article 6. <https://doi.org/10.3390/ijgi15010006>
- [18] 刘婧. 河北省地下水监测站网建设与成果应用的有关思考[J]. 河北水利, 2025(5): 14+30.
- [19] 郭天辰, 钱睿智, 刘海婧, 等. 地下水位对河水位的响应暨时序预测分析[J]. 陕西水利, 2025(5): 22-24.
- [20] Kanito, D., Benaafi, M. and Baalousha, H.M. (2025) Machine Learning Models for Groundwater Level Prediction and Uncertainty Analysis in Ruataniwha Basin, New Zealand. *Hydrology*, **12**, Article 282. <https://doi.org/10.3390/hydrology12110282>