Frontiers in Environmental Research
ISSN: 2960-026X
DOI: https://doi.org/10.61784/fer3024

GROUNDWATER LEVEL FITTING OF MONITORING WELLS IN THE BAODING REGION BASED ON LONG SHORT-TERM MEMORY (LSTM) NEURAL NETWORKS

Wei Guo
Institute of Hydrogeology and Environmental Geology, Chinese Academy of Geosciences, Shijiazhuang 5050061, Hebei, China.
Corresponding Email: 19831600205@163.com

Abstract: This study applies a Long Short-Term Memory (LSTM) neural network to model daily groundwater level variations at four monitoring wells in the piedmont plain of Baoding City, Hebei Province, China. Using daily data from January 1, 2018 to August 31, 2019, the LSTM model is trained and tested under a one-step rolling prediction framework. The results demonstrate that the LSTM model accurately fits groundwater time series across various hydrogeological conditions, with testing-phase RMSE values ranging from 0.08 to 0.45 meters and R^2 values exceeding 0.90. The model performs exceptionally well for both stable and fluctuating groundwater level conditions, capturing seasonal decline and recovery patterns without requiring explicit seasonal indicators. It also reveals the model's ability to learn long-term trends and nonlinear dynamics inherent in groundwater systems. Despite its high short-term prediction accuracy, challenges remain regarding multi-step forecasting and responses to extreme events outside the training data. The study concludes that LSTM offers strong potential for groundwater simulation in data-limited environments and recommends further integration of hydro-meteorological variables for enhanced model robustness.
Keywords: Groundwater level; LSTM; Time series modeling; Baoding; Deep learning

1 INTRODUCTION

Groundwater is a vital fundamental natural resource, playing a critical role in global water supply and agricultural production. As of 2010, the annual global extraction of groundwater exceeded 900 billion cubic meters, supplying approximately 36% of drinking water, 42% of irrigation water, and 24% of industrial water. In water-scarce regions of northern China, groundwater serves as the principal water source for both urban and agricultural use[1], earning the title "lifeline beneath the surface." However, because groundwater is stored underground in complex aquifer systems, its level dynamics are subject to various disturbances such as precipitation, evaporation, and pumping. The most direct way to observe groundwater level fluctuations is through long-term monitoring wells, yet the installation of such wells requires significant financial and human resources. Moreover, the number of wells is limited, their spatial distribution is uneven, and their performance is often constrained by geomorphic and topographic conditions, leading to missing data and outliers.

Against this background, machine learning techniques have increasingly been adopted in regional groundwater studies to make better use of sparse monitoring data. Since the 1990s, these methods have been widely applied in hydrological modeling and have demonstrated comparable or superior performance to numerical simulation models[2-4]. In recent years, recurrent neural networks (RNNs) and their variants—particularly the Long Short-Term Memory (LSTM) model—have shown significant promise in time-series groundwater level modeling and prediction [5-6]. LSTM networks learn temporal patterns automatically and are capable of capturing complex nonlinear relationships between input variables and groundwater level responses without requiring explicit assumptions about physical processes [7].

Studies have shown that LSTM generally outperforms traditional statistical and machine learning methods. For instance, Yin compared LSTM with Random Forest (RF) and Artificial Neural Networks (ANNs) for groundwater level prediction and found that LSTM achieved better accuracy in both validation and forecasting stages [8]. Vu demonstrated the model's ability to reconstruct groundwater time series with missing data and to accurately predict future values [9]. In China, Yan developed a multivariable LSTM model based on groundwater monitoring data from Tai'an, Shandong Province, integrating groundwater levels with various meteorological and anthropogenic factors. Their results demonstrated that the multivariable LSTM outperformed both the Backpropagation (BP) neural network and the univariate LSTM model, providing a more accurate simulation of groundwater dynamics[10]. Subsequently, Sun further optimized the model structure and input data processing in a case study of Jinan City. They found that fitting temperature series using sinusoidal functions and tuning the dropout rate significantly improved the model's stability and prediction accuracy. For the Quaternary aquifer, the optimized model achieved a prediction RMSE of less than 0.84 m[11]. Additionally, Sun compared ARIMA, BP neural networks, and LSTM for monthly and daily groundwater level prediction in the North China Plain, and confirmed the superiority of deep learning models in terms of accuracy [12].

Although the use of LSTM in groundwater modeling has become increasingly widespread, its application in fine-scale fitting of groundwater level time series in the piedmont plains of Baoding, Hebei Province, remains limited. Given the critical role of groundwater in regional water supply and the complexity of its dynamics in this area, this study aims to construct an LSTM-based model using daily groundwater level data from January 1, 2018 to August 31, 2019. The model will be trained and validated to evaluate its ability to simulate groundwater fluctuations and to assess its prediction accuracy and applicability, thereby offering insights for regional groundwater resource management..

2 STUDY AREA AND DATE

2.1 Study Area Overview

Baoding City is located in the northwestern part of the North China Plain. The piedmont zone of the Taihang Mountains refers to the transitional area between the eastern foothills of the Taihang Mountains and the adjoining plain, representing a region where the mountainous terrain of western Baoding gradually shifts eastward into the plains. The terrain in this region gradually flattens from west to east. Geologically, the area lies at the front edge of the alluvial-proluvial fan of the Taihang Mountains, consisting of thick Quaternary unconsolidated sediments that form the primary aquifer systems. Historically, this area was recognized as one of the most severely over-exploited groundwater zones in the North China Plain. Groundwater levels continuously declined since the late 20th century, forming a widespread groundwater depression cone, which reached its lowest point around 2010. In recent years, following the implementation of the South-to-North Water Diversion Project (Middle Route) and comprehensive measures to manage groundwater over-extraction [13], groundwater levels in the region have shown signs of recovery. Nevertheless, groundwater dynamics in this area remain highly sensitive and complex, calling for scientifically robust monitoring and modeling tools to support sustainable groundwater management.

2.2 Data Source

This study utilizes daily groundwater level elevation data from four national monitoring wells located in the piedmont plain of Baoding City, spanning the period from January 1, 2018 to August 31, 2019. The spatial distribution of these wells is illustrated in Figure 1. The selected wells cover a range of hydrogeological settings along a west–east transect across the piedmont plain, including both recharge zones near the Taihang Mountain foothills and overexploited zones within the central plain. The original time series is complete without any missing records. For a few isolated wells exhibiting abnormal peak values, such anomalies were identified as instrument errors and corrected through preprocessing: the outlier values were replaced by the average of the adjacent valid values to minimize the impact of outliers on model training. After preprocessing, the groundwater level time series for each monitoring well displayed a clear trend and are considered suitable for subsequent model training and validation.

Figure 1 Monitoring Well Locations（negative from google earth）
Source: https://earth.google.com/

3 METHODOLOY

3.1 Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) neural network is a variant of the recurrent neural network (RNN) architecture designed to handle long-term dependencies in sequential data[14]. Unlike traditional RNNs that are prone to vanishing or exploding gradient issues, LSTM introduces a gating mechanism that effectively controls the processes of information retention and forgetting. A standard LSTM unit consists of an input gate, a forget gate, an output gate, and a cell state. At each time step, the LSTM unit receives the current input vector, along with the previous hidden state and cell state, and computes the updated states through gate operations. The forget gate determines which information from the previous cell state should be discarded based on the current input and previous hidden state. The input gate regulates the extent to which new information is written into the cell state. The updated cell state combines these two processes to maintain long-term memory. Finally, the output gate determines which parts of the cell state are passed to the hidden state output. These gating operations enable LSTM units to preserve relevant information over tens or even hundreds of time steps, making the model well-suited for capturing long-term trends and seasonal patterns in groundwater level time series[15]. Mathematically, the LSTM computations involve sigmoid activations for the gates and tanh activation for candidate memory states; the detailed equations are omitted here for brevity. The model is trained using gradient descent to update the gate weights, minimizing the difference between the predicted and observed sequences. Compared to traditional time-series methods or shallow neural networks, LSTM's ability to retain long-range dependencies provides a significant advantage in modeling delayed-response processes such as groundwater level dynamics[16].

3.2 Model Architecture and Parameter Configuration

The LSTM model was implemented and trained using Python. For each monitoring well, the groundwater level time series was transformed into supervised learning data by creating input-output pairs. Given the relatively stable daily changes and seasonal cycles of groundwater levels, a sliding window of $N = 30$ days (approximately one month) was used as the input sequence length. Each input sequence consisted of groundwater levels over the previous 30 days, and the target output was the groundwater level on the following day ($N+1$). This sliding window approach was applied to extract training samples from the historical time series. To ensure data consistency across wells, the raw groundwater levels were normalized using Min-Max scaling, mapping values to the range [0, 1]. After prediction, the outputs were rescaled back to their original range via inverse normalization.

The LSTM model architecture comprised a single LSTM hidden layer followed by a fully connected output layer. The number of hidden units was determined experimentally to balance accuracy and prevent overfitting, and 16 units were selected. The hidden layer used the tanh activation function, while the output layer employed a linear activation to generate continuous numerical predictions. The model was trained using the Backpropagation Through Time (BPTT) algorithm, with the Adam optimizer and an initial learning rate of 0.01, which was adaptively adjusted during training to accelerate convergence. The loss function used was Mean Squared Error (MSE). To improve generalization, the training data was randomly shuffled, and mini-batch gradient descent (batch size = 16) was applied for parameter updates. The number of training epochs was determined based on convergence trends in the validation loss, and training was terminated when validation error ceased to decrease.

3.3 Training and Testing Set Division

To evaluate the generalization performance of the LSTM model, the complete time series was split into training and testing subsets. Following a conventional 80/20 split, the period from January 1, 2018 to April 30, 2019 (approximately 486 days) was used as the training set, while the period from May 1, 2019 to August 31, 2019 (approximately 122 days) was used as the independent test set. It is important to note that model training was conducted exclusively on the training data, and the testing phase incorporated a rolling prediction mechanism. For example, to predict groundwater levels on May 2, 2019, the model used real observations from the 30-day period ending on May 1 as input. The model then predicted the next day, and this process continued iteratively using actual observed values to update the input window, ensuring sequential coherence in multi-step forecasting.

3.4 Evaluation Metrics

The model's prediction performance was evaluated on the test set using two common metrics: Root Mean Square Error (RMSE) and the Coefficient of Determination (R^2). RMSE quantifies the average deviation between predicted and observed values, with lower values indicating better accuracy. R^2 measures the proportion of variance in the observed data explained by the model, with values closer to 1 indicating better fit. These metrics were used to quantitatively assess the LSTM's performance across different monitoring wells. In addition, predicted and observed groundwater level curves were plotted for visual inspection, allowing for a direct comparison of the model's simulation capability and error distribution.

4 MODEL FITTING AND RESULT ANALYSIS

4.1 Model Training Process

Following the methodology described above, separate LSTM models were trained for the groundwater level time series of the four monitoring wells. During the training phase, the loss function for each model exhibited a rapid decline followed by stabilization, indicating effective convergence. After approximately 100 epochs, the training error reached a plateau, and the validation error no longer decreased, suggesting that the model had achieved its optimal fitting capacity. At this point, the Root Mean Square Error (RMSE) on the training set for each well ranged between 0.1 and 0.3 meters, demonstrating a high level of fitting accuracy on the observed data (Figure 2).

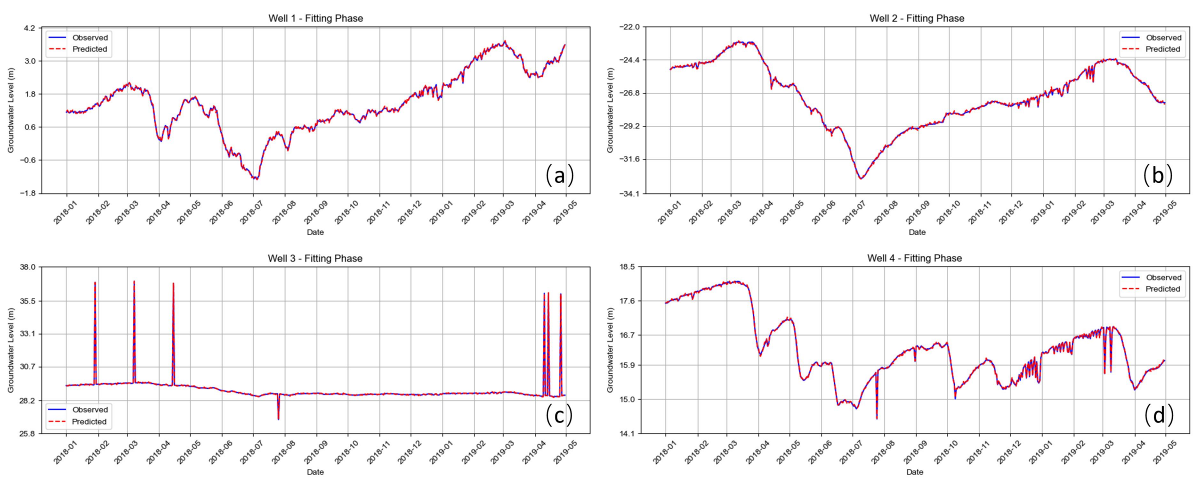

Figure 2 Groundwater Level Fitting on the Training Set. Subplots (a)–(d) Correspond to Wells 1 through 4, Respectively. The Red Dashed Lines Represent the Predicted Groundwater Levels, while the Blue Solid Lines Indicate Observed Values.

4.2 Testing Phase Fitting Performance

The trained models were then applied to the testing set, which spanned from May to August 2019, to generate daily predictions of groundwater levels. The results are shown in Figure 3, where the predicted groundwater levels are compared against observed values for each of the four monitoring wells. The LSTM models effectively reproduced the temporal dynamics of groundwater levels, with strong agreement between the predicted and observed curves. In terms of seasonal variation, the models successfully captured the summer decline and early autumn recovery patterns typical of groundwater levels. Furthermore, the models accurately reflected broader trends of sustained decline or gradual rise in groundwater levels across different wells.

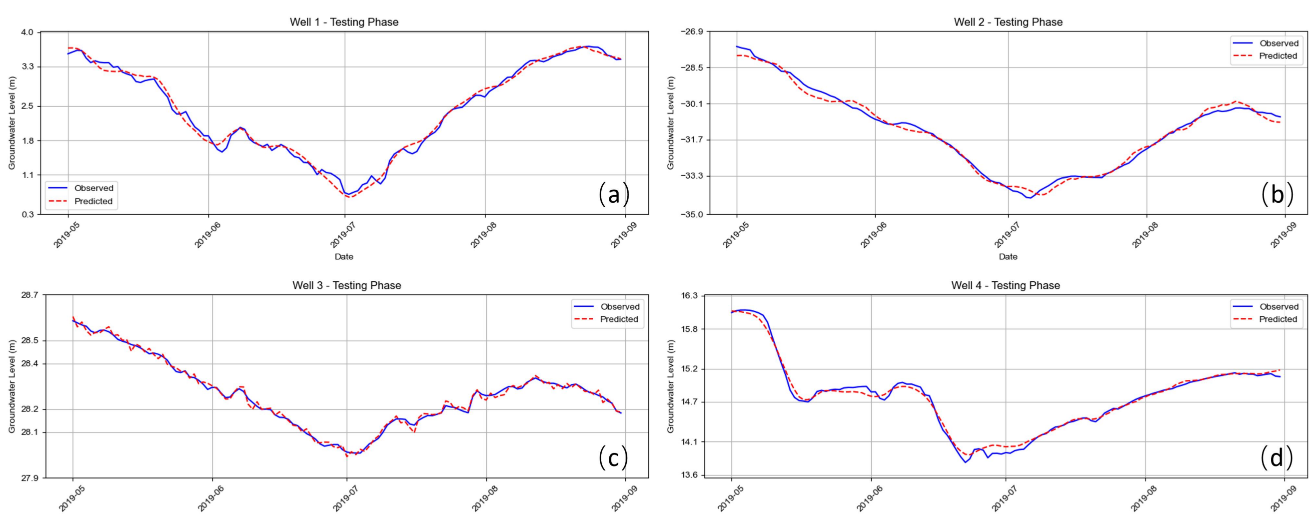

Figure 3 Groundwater Level Fitting on the Testing Set. Subplots (a)–(d) Correspond to Wells 1 through 4, Respectively. The Red Dashed Lines Represent LSTM-predicted Groundwater Levels, and the Blue Solid Lines Denote Actual Observed Values.

4.3 Error Analysis

The four monitoring wells selected for this study exhibited distinct groundwater level variation patterns during the testing period from May to August 2019, yet the LSTM model demonstrated satisfactory fitting performance across all cases (Table 1). For Well 1, the groundwater level first increased and then decreased, dropping from approximately 3.2 m to 1.8 m. The LSTM model successfully captured both the overall trend and the inflection points, with an RMSE of 0.21 m and an R^2 of 0.94. Well 2 is located in a region of intensive groundwater overexploitation and experienced a sharp decline of nearly 5 m. The model accurately reproduced the nonlinear trend and slope transitions, although a slight underestimation occurred toward the end of the period; the RMSE and R^2 were 0.45 m and 0.90, respectively. Well 3 exhibited stable water levels with a total decline of less than 1 m. The model's predictions were nearly indistinguishable from the observed data, achieving an RMSE of only 0.08 m and a high R^2 of 0.99, indicating excellent fitting accuracy. For Well 4, the groundwater level dropped by approximately 1.3 m, accompanied by minor fluctuations. The LSTM model accurately replicated both the overall trend and localized short-term rebounds, with an RMSE of 0.18 m and R^2 of 0.96.

Table 1 Monitoring Well Error Statistics

Monitoring well number	RMSE (training set)	R^2 (training set)	RMSE (test set)	R^2 (test set)
Well 1	0.036	0.99	0.21	0.94
Well 2	0.073	0.99	0.45	0.9
Well 3	0.028	0.99	0.08	0.99
Well 4	0.026	0.99	0.18	0.96

In summary, the LSTM model demonstrated strong nonlinear learning capability and high accuracy across varying hydrogeological settings and magnitudes of groundwater level change. It effectively captured the dynamic characteristics of groundwater fluctuations in the study area, confirming its applicability in regional groundwater modeling.

5 CONCLUSIONS AND FUTURE PERSPECTIVES

This study applied a Long Short-Term Memory (LSTM) neural network model to simulate daily groundwater level variations at four monitoring wells in the piedmont plain of Baoding City. The main conclusions are as follows:

The LSTM model achieved high-precision fitting of groundwater level time series. After being trained on approximately 1.5 years of daily data, the model produced predictions that closely matched observed groundwater levels during the testing phase. The mean absolute error ranged from 0.1 to 0.45 meters, and the coefficient of determination (R^2) generally exceeded 0.90, indicating that the model could explain the majority of groundwater level variance. Particularly for wells with stable and gradually changing levels, the fitting accuracy approached the magnitude of measurement error. For wells with significant seasonal fluctuations or trends, the LSTM model effectively captured the peaks, troughs, and long-term trajectories. These results confirm that LSTM networks are well-suited for modeling hydrological time series such as groundwater levels, which exhibit long-term memory characteristics.

The model automatically learned seasonal and trend components. Leveraging the memory capabilities of gated recurrent units, the LSTM model captured seasonal patterns—such as summer declines and winter recoveries—as well as long-term trends like multi-year depletion, without requiring explicit inclusion of seasonal or trend indicators. During the testing period, the model responded accurately to seasonal turning points, distinguishing dry and wet periods. This demonstrates the advantage of deep learning models in capturing latent patterns. In contrast, traditional approaches like linear regression or ARIMA typically require pre-processing (e.g., detrending or adding exogenous seasonal terms), whereas LSTM can simultaneously model multiple forms of variation.

While the model performed exceptionally well in short-term predictions, its accuracy over longer forecasting horizons remains uncertain. This study focused on one-step-ahead rolling forecasts, using recent observed values to update predictions continuously. Under these conditions, the model achieved excellent results (e.g., $R^2 = 0.99$). However, for multi-step forecasts extending weeks or months into the future, prediction errors may accumulate over time. In such cases, periodic correction or hybridization with physically-based models may be needed to mitigate error propagation.

The model's performance is dependent on the range of training data and may have limited adaptability to abnormal or abrupt conditions. As the LSTM model learns entirely from historical data, it may struggle to respond accurately to events not represented in the training set, such as sudden recharge from extreme rainfall or rapid drawdown from unexpected over-pumping. In this study, abnormal values were filtered to improve model stability. Nonetheless, this highlights the need to supplement model outputs with real-time field data to validate and adjust predictions, particularly when encountering conditions outside the historical data distribution.

In summary, the LSTM model delivered satisfactory results for groundwater level fitting in the Baoding piedmont plain, demonstrating strong potential as a data-driven modeling tool. Nevertheless, there remain areas for improvement and further investigation. For instance, this study used only historical groundwater level data as model input. In future work, incorporating additional variables such as precipitation, evaporation, and pumping rates may help build a multivariate predictive model with enhanced robustness and improved responsiveness to anomalous scenarios.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Wang K, Chen H, Fu S, et al. Analysis of exploitation control in typical groundwater over-exploited area in North China Plain. Hydrological Sciences Journal, 2021, 66(5): 851-861. DOI: 10.1080/02626667.2021.1900575.
[2] Kratzert F, Klotz D, Brenner C, et al. Rainfall–runoff modelling using long short-term memory (LSTM) networks. Hydrology and Earth System Sciences, 2018, 22(11): 6005-6022. DOI: 10.5194/hess-2018-247-sc1.
[3] Lees T, Buechel M, Anderson B, et al. Benchmarking data-driven rainfall–runoff models in Great Britain: a comparison of long short-term memory (LSTM)-based models with four lumped conceptual models. Hydrology and Earth System Sciences, 2021, 25(10): 5517-5534. DOI: 10.5194/hess-25-5517-2021.
[4] Yang S, Yang D, Chen J, et al. A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. Journal of Hydrology, 2020, 590: 125206. DOI: 10.1016/j.jhydrol.2020.125206.
[5] Zounemat-Kermani M, Batelaan O, Fadaee M, et al. Ensemble machine learning paradigms in hydrology: A review. Journal of Hydrology, 2021, 598: 126266. DOI: 10.1016/j.jhydrol.2021.126266.
[6] Nearing G S, Kratzert F, Sampson A K, et al. What role does hydrological science play in the age of machine learning?. Water Resources Research, 2021, 57(3): e2020WR028091. DOI: 10.31223/osf.io/3sx6g.
[7] Bai T, Tahmasebi P. Graph neural network for groundwater level forecasting. Journal of Hydrology, 2023, 616: 128792. DOI: 10.1016/j.jhydrol.2022.128792.
[8] Yin W, Fan Z, Tangdamrongsub N, et al. Comparison of physical and data-driven models to forecast groundwater level changes with the inclusion of GRACE–A case study over the state of Victoria, Australia. Journal of Hydrology, 2021, 602: 126735. DOI: 10.1016/j.jhydrol.2021.126735.
[9] Vu T D, Ni C F, Li W C, et al. Predictions of groundwater vulnerability and sustainability by an integrated index-overlay method and physical-based numerical model. Journal of Hydrology, 2021, 596: 126082. DOI: 10.1016/j.jhydrol.2021.126082.
[10] Baizhong Y, Jian S, Xinzhou W, et al. Multivariable LSTM neural network model for groundwater levels prediction.Journal of Jilin University (Earth Science Edition), 2020, 50(1): 208-216. DOI: 10.13278/j.cnki.jjuese.20190055.
[11] Hongjie S, Zhenhua Z, Linxian H, et al. Application of Multi-Variable LSTM Neural Network Model for Groundwater Levels Prediction. Yellow River, 2022, 44(8). DOI: 10.3969 /j.issn.1000-1379.2022.08.014.
[12] Sun J, Hu L, Li D, et al. Data-driven models for accurate groundwater level prediction and their practical significance in groundwater management. Journal of Hydrology, 2022, 608: 127630. DOI: 10.1016/j.jhydrol.2022.127630.
[13] Zhang C, Duan Q, Yeh P J F, et al. Sub-regional groundwater storage recovery in North China Plain after the South-to-North water diversion project. Journal of Hydrology, 2021, 597. DOI: 10.1016/j.jhydrol.2021.126156.
[14] Rajaee T, Ebrahimi H, Nourani V. A review of the artificial intelligence methods in groundwater level modeling. Journal of hydrology, 2019, 572, 336-351. DOI: 10.21203/rs.3.rs-2915223/v1.
[15] Wai K P, Chia M Y, Koo C H, et al. Applications of deep learning in water quality management: A state-of-the-art review. Journal of Hydrology, 2022,613, 128332. DOI: 10.1016/j.jhydrol.2022.128332.
[16] Van Thieu N, Barma S D, Van Lam T, et al. Groundwater level modeling using augmented artificial ecosystem optimization. Journal of Hydrology,2023, 617, 129034. DOI: 10.1016/j.jhydrol.2022.129034.