Public Health and Environment 2025, 1(2), 41–60.
https://doi.org/10.70737/jvxmvd71

Article

Predicting Surface Water Level based on Wavelet Transform and Long Short-term Memory Network

Beiyan Li [1], Junbo Yu [1], Jin Zhang [2], Cong Liu [3], Yan Zhao [4], Yuanxin Liu [1], Linyi Guo [1], Lu Lv [1], Rongqi Liu [1], Weiwei Yu [1, *]

[1] Key Laboratory of Hydraulic and Waterway Engineering of the Ministry of Education, College of River and Ocean Engineering, Chongqing Jiaotong University, Chongqing 400074, China
[2] School of Geography, South China Normal University, Guangzhou 510631, China
[3] Southwest Institute of Technology and Engineering, Chongqing 400039, China
[4] Shanghai Municipal Engineering Design Institute (Group) Co. Ltd., Shanghai 200003, China
*** Corresponding author:** yu11237@cqjtu.edu.cn

CITATION

Li B., Yu J., Zhang J., Liu C., Zhao Y., Liu Y., Guo L., Lv L., Liu R., Yu W. Predicting Surface Water Level based on Wavelet Transform and Long Short-term Memory Network. Public Health and Environment. 2025, 1(2): 41–60. https://doi.org/10.70737/jvxmvd71

ARTICLE INFO

Received: 27 August 2025
Accepted: 12 November 2025
Available online: 3 December 2025

Abstract: Water level prediction work is crucial to provide timely water information for flood control and drought relief and shipping. In this study, a hybrid surface water level prediction model (W-LSTM) based on wavelet transform and deep learning model LSTM was constructed to provide an accurate surface water level modeling prediction method. First, to extract the hidden features behind the original water level time series, it is decomposed into multiple sub-series utilizing wavelet transform. Then, combine with a deep learning algorithm to build an LSTM model for each component to make the prediction. Finally, the predicted values of the individual components are summed to produce the ultimate prediction results. The results show that after pre-processing by wavelet transform and then combining the advantages of the LSTM algorithm can significantly improve the prediction accuracy, and W-LSTM demonstrates better prediction power than the other comparative models (BPNN, ELM, LSTM, W-NN, and W-ELM). The evaluation criteria R^2, *RMSE*, and *MAE* of the model are 0.9575, 0.1465, and 0.0737, respectively. The W-LSTM model can achieve accurate and effective water level prediction, which can provide a scientific reference basis for water environment management.

Keywords: wavelet transform; LSTM; time series; water level prediction; hybrid model

1. Introduction

In recent years, extreme climate change has led to frequent intense rainstorms, thus significantly increasing the likelihood of river flooding [1]. Flooding from heavy rainfall has caused great economic losses and human casualties [2,3] and has led to an increase in the frequency and severity of this risk year by year with the increase in extreme weather events [4]. According to relevant statistics, floods caused nearly 70,000 fatalities worldwide in the 20th century, while floods have accounted for approximately 40% of global natural disaster losses from 1980 to the present, totaling over US$1 trillion [5,6]. Therefore, it has become imperative to optimize and strengthen flood protection measures. Water level prediction, as one of the countermeasures, has now received wide attention and has been successfully applied in rivers with flood risk [7]. In addition, water level changes are an important factor in the waterway transportation process [8] and a key point of primary consideration in the shipping process. Given this, accurate water level prediction can not only help flood control and reduce the damage caused by floods [9], but also facilitate the

maintenance of waterways and implement appropriate early warning programs for the relevant decision-makers, thereby enhancing traffic safety and navigability [10].

In recent decades, several modeling methods have been proposed for water level predictions. In the past, traditional water level prediction methods were mainly based on physical and process-based models [11,12], including grid-based distributed hydrological modeling [13], Muskingum model parameters optimal estimation method [14], least-squares-based harmonic analysis method [15], advanced hydrologic prediction system (AHPS) [16], and generalized extreme value distribution model (GEV) [17]. Although these methods can build corresponding water level prediction models based on hydrology and simulate changes in water level, they require the use of extensive hydrological data [18,19]. In practice, however, much of the required field data is often unavailable, which limits the performance of physically based models and thus impedes accurate real-time or fixed-point predictions [11]. In addition, process-based models are generally time-consuming [20], which is a significant disadvantage in practical applications. Therefore, considering the long-term variability of water levels, these modeling methods are often not practical for operational use.

As both theory and practice advance, machine learning (ML) inspired artificial neural network (ANN) techniques have received widespread attention due to their high fault tolerance, high prediction accuracy, and reliability in dealing with complex non-linear relationships between variables [21]. These methods have been successfully applied to water level prediction. These methods include radial basis function neural network (RBF) [22], extreme learning machine (ELM) [23], back propagation neural network (BPNN) [24], and recurrent neural network (RNN) [25]. Among these methods, the long short-term memory network (LSTM) optimizes the shortcomings in RNN, takes into account the long-term dependency problem existing in the time series, and is capable of capturing the long-term time correlations in the series, which effectively enhances the accuracy of the prediction [26]. Compared to RNN, LSTM can capture information from data points further back in time, and it has demonstrated superior performance, especially in longer time series, in water level prediction [27]. Although ANN is effective in many fields of prediction, it still has some limitations and shortcomings when dealing with non-stationary data sets [28].

Hydrological processes in nature are influenced by a variety of physical factors, including rainfall characteristics, topographic features, and climatic conditions, which lead to complex variability. This results in hydrological time series exhibiting non-linear, non-stationary, and multi-scale characteristics [29,30]. As an effective tool for analyzing non-stationary data sets containing multi-scale features, wavelet transform (WT) has proven that in the process of analyzing hydrological time series, various features hidden behind the data structure can be deeply excavated, thus revealing the correlation and coherence of hydrological time series and helping to improve the modeling ability [31,32]. The use of wavelet transform to process original time series allows for the identification of trends at different time scales, and has become increasingly popular in recent years for data-driven hydrological modelling [33].

Therefore, to effectively manage the water environment while improving the prediction performance of modeling, this study proposes a surface water level prediction method (W-LSTM) based on wavelet transform and the deep learning

model LSTM. The focuses of this study were: (1) decomposition of water level time series using wavelet transform, which aids in exploring and extractionthe internal features of the time series, (2) to develop the W-LSTM hybrid water level prediction model and provide comprehensive prediction of water level for the study area, and (3) to verification the performance of the W-LSTM model through comparison with several other models and evaluation criteria.

2. Materials and Methods

2.1. Study Area and Water Level Data

Le'an River is situated in the northeast of Jiangxi Province, China (116°35′~118°03′E, 29°03′~29°34′N), which is the upper mainstream of Rao River, a tributary of Poyang Lake in Yangtze River Basin, and crosses Wuyuan, Dexing, Leping, Wannian, and Poyang counties and cities, and is the main water resource of Shangrao area and an important navigation channel in the basin. There are many mountain ranges in the basin of the Le'an River, and the topography and natural geographical conditions in the basin are favorable to the formation of storm floods, which makes the basin very prone to storm flood processes. Therefore, it is important to establish an effective and accurate water level forecasting system to provide a scientific basis for flood control decisions in the basin and to safeguard and improve navigability in the basin. The daily historical water level data monitored by the Sandu hydrological station (Figure 1) in the Le'an River basin were used for this study, which was monitored from January 1, 2019 to June 30, 2022 at 8:00 am every day, from a total of 1277 data sets. The recorded data was divided into a training set (80%) and a test set (20%). Because the water-quality series are typically nonlinear and large-scale, standard outlier heuristics (Z-score, box-plot, visual screening) are limited by distributional assumptions, scale effects, and subjectivity, while conventional imputation (linear or spline interpolation) can be unreliable—linear methods presuppose local linearity, and splines are better suited to small gaps. Accordingly, we removed outliers using the Pauta criterion (3σ rule) and imputed missing values via k-nearest neighbors (KNN). The model is trained using the training set and the performance of the constructed model is verified using the test set [34].

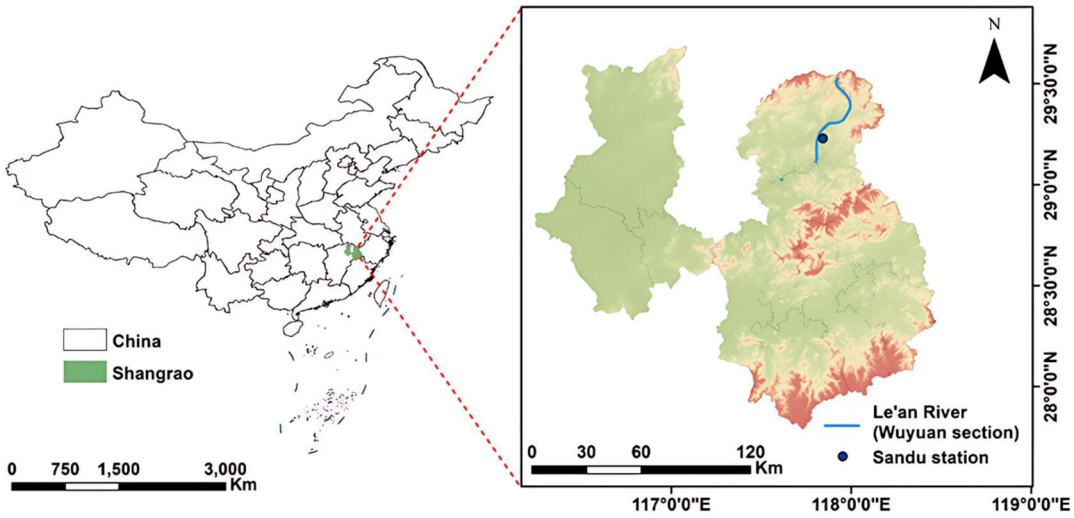

Figure 1. Location of the hydrological monitoring station.

2.2. Simulation Experiment Environment

All model algorithms constructed and utilized in this study were implemented using MATLAB R2019b (MathWorks, Natick, MA, USA) programming software. The specific simulation experimental environment is shown in Table 1.

Table 1. Simulation Experiment Environment Configuration.

Equipment	Title 2
System	Windows10 64bit
Processor	Intel(R) Core(TM) i5-7300HQ CPU @ 2.50GHz
RAM	8G

2.3. Parameter Settings

To ensure rapid convergence while maintaining accuracy and avoiding overfitting, this study employs multiple trial methods to optimally configure key hyperparameters. Primary considerations include the number of hidden units, the number of intermediate layers, the number of iterations, learning rate, batch size, and Dropout ratio. The sliding window size was set to 1 (single-step lookahead), and the Adam optimizer was employed. The hidden layer employs standard LSTM cells (sigmoid gates, tanh state updates), while the output layer uses a linear connection to regress to the original scale. A non-recurrent Dropout of 0.20 is applied after each LSTM layer, combined with early stopping and gradient clipping.

2.4. Wavelet Transform Theory

The wavelet transform is a special mathematical transform that is a further development of the Fourier transform [35]. The Fourier transform is a common analysis method in the field of digital signals, however, it has a serious shortcoming

in that it discards the time information in the transform. In contrast, the wavelet transform overcomes the shortcoming that the Fourier transform is localized only in frequency, and is a transform analysis tool with both time and frequency domains. By analyzing the original time series using the wavelet transform, the time and frequency information in the time series can be extracted.

A wavelet basis function is the foundation of the wavelet transform. According to the continuity of wavelet basis function changes, wavelet transform can be subdivided into discrete wavelet transform (DWT) and continuous wavelet transform (CWT). The CWT can be used to reveal the characteristics of series at multiple time scales, so in this paper, we first use CWT to perform a multi-scale analysis of the water level time series, which is defined as

$$W(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t)\,\overline{\psi}\left(\frac{t-b}{a}\right) dt \tag{1}$$

where $f(t)$ is the time series, a and b control the scale and translation of $f(t)$, respectively; ψ is the wavelet function and satisfies $\int_{-\infty}^{+\infty} \psi(t)\,dt = 0$, and $\bar{\psi}$ is the conjugate complex of ψ.

The DWT is a signal decomposition method [36]. Unlike the CWT, the DWT transforms only certain scales in the process of transforming, rather than continuous transform parameters. The time series can be decomposed into multiple sub-series by using discrete wavelet transform, so that the time series can achieve multi-resolution refinement, which is defined as

$$W_f(j,k) = |a_0|^{-\frac{j}{2}} \sum_{t=0}^{n-1} f(t)\overline{\psi}({a_0}^{-j}t - kb_0) \tag{2}$$

where a_0 and b_0 are constants, j is the decomposition level, k is the translation parameter, and n is the signal length.

Considering the relatively complex environment of water bodies in nature, the observed hydrological series are usually discrete signals, so most hydrological prediction problems give priority to the discrete wavelet transform. Therefore, this study further decomposes the water level time series by using DWT. The decomposition of the signal is the process of filtering, which is achieved by utilizing filters (high-pass and low-pass). The signal passing through the filter results in two sub-series: the detail component D (high frequency) and the approximation component A (low frequency). Figure 2 shows a schematic of the DWT decomposition. The approximate component A for the low frequencies can be further decomposed until the maximum decomposition level is reached. The final structure after passing through n levels of decomposition is [A_n, D_n, D_{n-1}, ..., D_2, D_1] [37]. The low-frequency series reflects the main variability features of the time series, and the high-frequency series reflects the irregular fluctuations in the time series caused by uncertain influences.

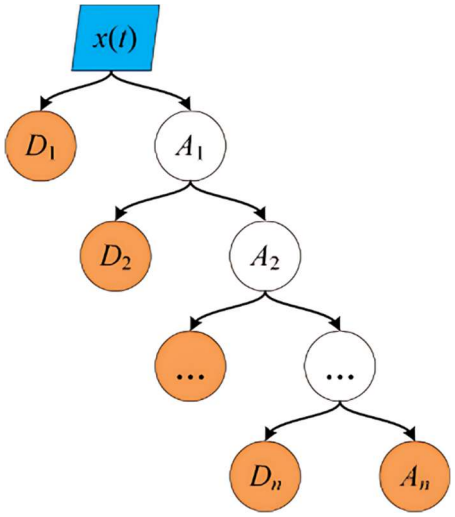

Figure 2. DWT decomposition schematic.

After the wavelet transform is completed, the low-frequency and high-frequency series obtained by decomposition are inversely transformed, namely, wavelet reconstruction. The series obtained after reconstruction is the same length as the original series, and the original series can be regarded as the superposition of the reconstructed series, which is defined as

$$x(t) = A_n + \sum_{t=1}^{n} D_i \tag{3}$$

where A_n is the low-frequency reconstruction series at the maximum decomposition level, D_i is the high-frequency reconstruction series at each decomposition level.

2.5. Deep Learning Model LSTM

The long short-term memory (LSTM) network is a variant structure of the RNN model proposed by Hochreiter and Schmidhuber that avoids the long-term dependency problem by deliberate design [38]. RNN is an effective tool for processing time series, however, it has a serious drawback that RNN suffers from gradient disappearance and gradient explosion problems in training [39]. In contrast, LSTM effectively overcomes this shortcoming and allows better learning of long-time series without the gradient problem. Meanwhile, LSTM is a deep learning model with better feature representation and convergence ability for non-linear datasets compared to shallow learning models [40].

Figure 3 shows the structure of the LSTM, which adds four structures in the hidden layer compared to RNN: forget gate, input gate, cell state, and output gate [41]. LSTM controls the flow of information through gating units to prevent the interference that useless data may cause during the prediction process, allowing better exploitation of the time-varying properties of time series data. The specific details are as follows.

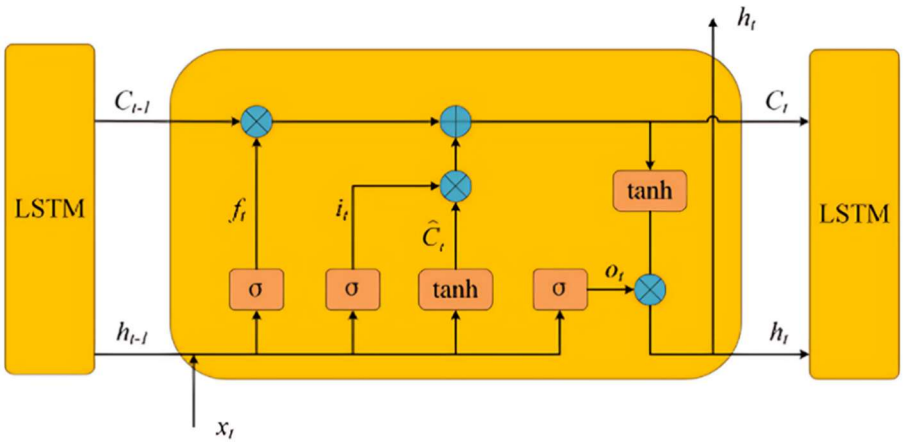

Figure 3. LSTM network structure diagram.

The forget gate f_t selectively retains the information transmitted from the previous cell state C_{t-1}. The activation function sigmoid determines how much information is retained by the f_t.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$

The input gate i_t selectively memorizes the input x_t of the current phase and the hidden state information h_{t-1} of the previous phase by means of a sigmoid function. The updated cell state C_t is determined jointly by the previous cell state C_{t-1} and the candidate cell state $\hat{C}_t$.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{5}$$

$$\hat{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{6}$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \hat{C}_t \tag{7}$$

Similarly, the output gate o_t controls the final output h_t transmitted to the next hidden layer through the sigmoid function.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{8}$$

$$h_t = o_t \otimes tanh(C_t) \tag{9}$$

In the above formulas, W_f and b_f, W_i and b_i, W_c and b_c, and W_o and b_o are the weight coefficients and bias values of the forget gate, input gate, cell state, and output gate, respectively; σ is the activation function sigmoid with a value between [0,1]; tanh is the hyperbolic tangent activation function.

2.6. Wavelet-LSTM Model (W-LSTM)

The water level time series are generally affected by various factors, such as water volume, rainfall and tides, and other natural factors and related human activities, showing characteristics such as uneven and non-linear. Therefore, a hybrid prediction model (W-LSTM) based on wavelet transform and LSTM is constructed in this study to achieve effective prediction of surface water level. By decomposing the original water level data through discrete wavelet transform and then combining the advantages of the LSTM model algorithm, more accurate and efficient prediction of

water level time series data can be achieved. The modeling process of the constructed W-LSTM is shown in Figure 4, and the specific steps are as follows:

Step 1. Decomposition of original water level data into multiple components using DWT, including n high-frequency components *D* and a low-frequency component *A*.

Step 2. The decomposed series have different lengths, so the n+1 components of the decomposition are wavelet reconstructed to obtain the n+1 sub-series so that they have the same length as the original series.

Step 3. To accelerate training and improve numerical stability, all variables were normalized to [0,1] using Equation. (10) prior to model fitting. We adopted per-component min–max scaling (parameters estimated on the training split only) because the resulting input range aligns with the LSTM's sigmoid/tanh gating, reducing activation saturation and stabilizing gradients, and because the transform is exactly invertible, which facilitates post-prediction inverse scaling and wavelet-component reconstruction. In preliminary comparisons, predictive performance under min–max was essentially indistinguishable from that obtained with z-score normalization; accordingly, we retained the simpler, invertible min–max scheme.

$$x_{Norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{10}$$

Step 4. The LSTM model is constructed separately for the n+1 sub-series obtained after decomposition, and the corresponding predictions are made. In the prediction process, the historical water level data of the T−1 moment and the T moment are used as input and output variables, respectively; that is, the water level of the next day is predicted using the water level data of the previous day.

Step 5. The prediction outputs are inversely normalized, and finally, the component predictions are summed to produce the final results of the water level prediction.

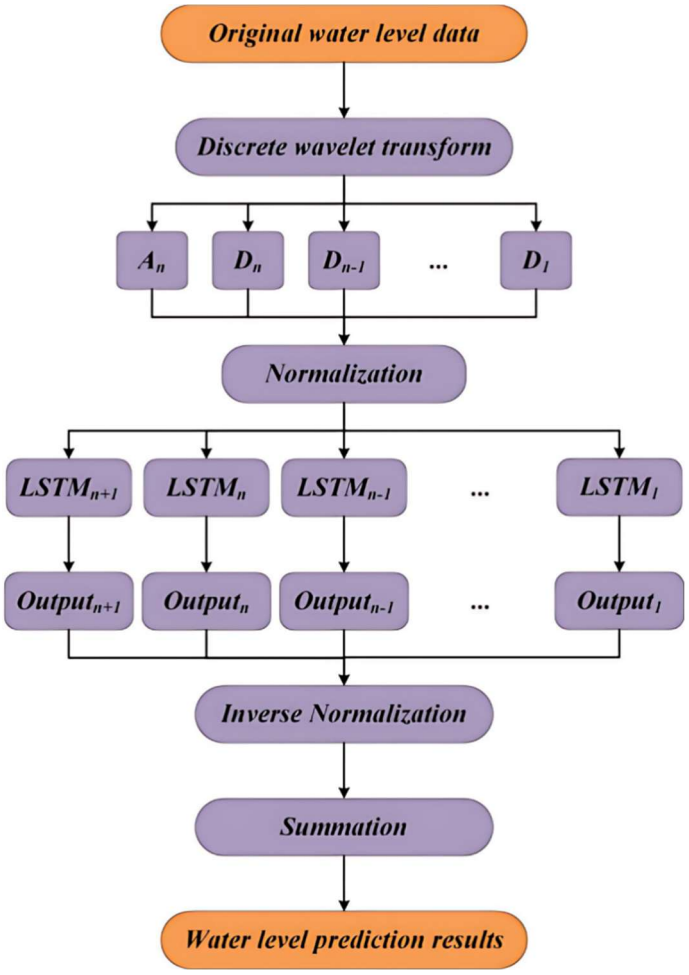

Figure 4. W-LSTM model prediction flow chart.

2.7. Performance Evaluation

Different evaluation criteria can be used to assess the stability and effectiveness of the model from several different aspects. Therefore, to quantify the predictive ability of the proposed method, this study utilizes the root mean square error (*RMSE*), mean absolute error (*MAE*) and coefficient of determination (R^2) to assess the performance of the model. Equation. (11), Equation. (12), and Equation. (13) are the calculation methods for the three evaluation criteria, respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{measured} - y_{predicted})^2} \tag{11}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{measured} - y_{predicted}\right| \tag{12}$$

$$R^2 = 1 - \frac{\sum \left(y_{measured} - y_{predicted}\right)^2}{\sum \left(y_{measured} - y_{mean-measured}\right)^2} \tag{13}$$

3. Results

Three classic conventional models (BPNN, ELM, and LSTM) and hybrid models combined with wavelet transform (W-NN, W-ELM) were chosen as the comparative

models to assess the performance of the developed model (W-LSTM) for surface water level prediction. Also, according to section 2.1, the first 1020 datasets are used for training the models, and the last 255 datasets are used for simulation prediction. All model algorithms in this study are implemented by MATLAB R2019b software.

3.1. Study Area Water Level Characteristics

The distribution of water level in Le'an River during the study period is shown in Figure 5(a), the distribution of the water level ranges from 52.97 to 60.51 m, with the average water level around 53.64 m. In addition, according to the histogram (Figure 5(b)), it can be noted that the water level during the study period was mainly concentrated between 53–55 m, and the number of its records showed a sharp decreasing trend with the increase of water level, thus reflecting that the overall distribution interval of water level in Le'an River was relatively stable. However, considering the Le'an River as a shipping waterway around the route and the uneven distribution of precipitation within the territory during the year, it is prone to heavy rainfall and thus highly susceptible to the formation of flood processes. Therefore, surface water level prediction for it still seems necessary.

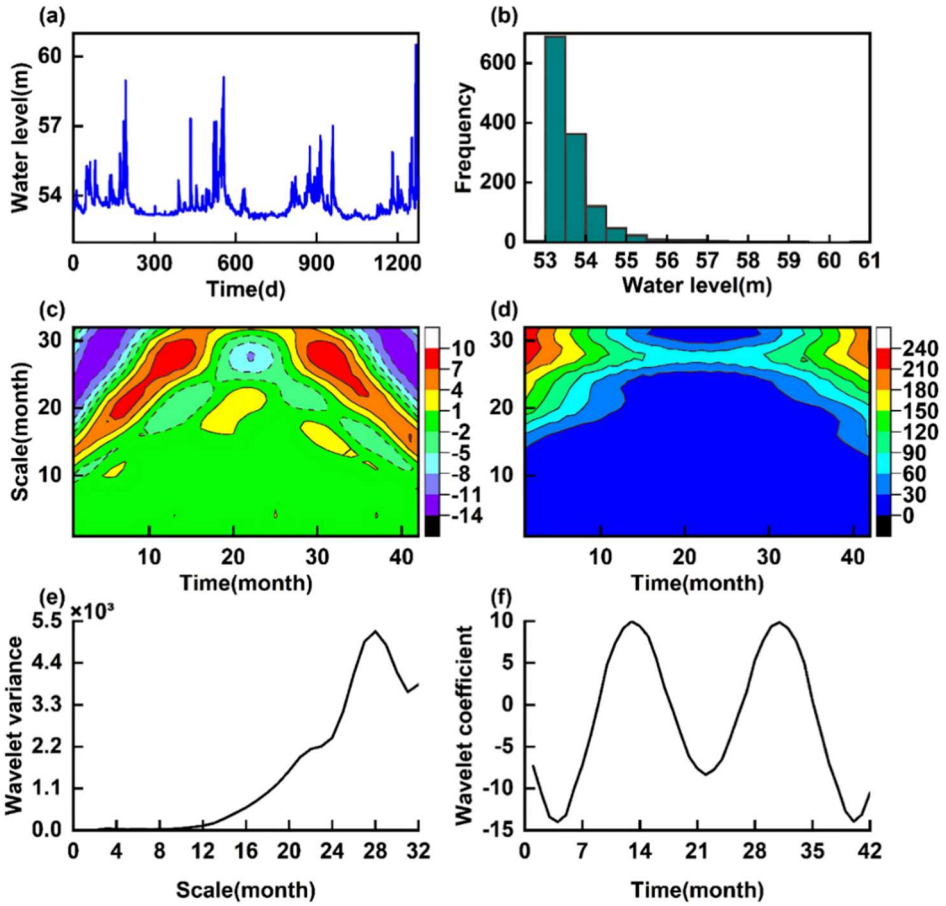

Figure 5. (**a**) Distribution plot, (**b**) histogram, (**c**) wavelet coefficient real part plot, (**d**) wavelet coefficient modulus squared plot, (**e**) wavelet variance plot, and (**f**) trend plot of the first main period of water level in Le'an River.

The periodic variations in the monthly mean water level during the study period were further investigated using CWT. Figure 5(c) and Figure 5(d) show the wavelet coefficient real part plot and wavelet coefficient modulus squared plot of the water level data in the study area, respectively. It is evident that the study period was dominated by oscillatory period on the 12–32 months scale, and that oscillation on the 18–30 month scale was more pronounced, with the oscillation center at 28 months. Four alternating changes in dry and abundant water levels occurred throughout the time scale, between months 4, 14, 22, 31, and 40, respectively. As can be noted in Figure 5(d), the 18–32 months scale has the strongest energy, indicating that the periodicity is most pronounced during this period. However, it is worth noting that there is a localization of the periodic variation during this period, before the 15th month and after the 29th month. Although the energy is weaker on the 14–18 months scale, the oscillatory period covers the whole study period.

Figure 5(e) shows the wavelet variance plot of the water level data, and there is only one peak at 28 months in the plot, indicating that 28 months is the first main period. Based on the corresponding main period trend plot (Figure 5(f)), it can be observed that on the 28 months' time scale, the period of water level changes is 18 months, and exhibits two periods of alternating changes between dry and abundant water levels. This series of periodic variations constitutes the multi-scale feature of the water level series in the region, causing challenges in predicting water levels accurately. Therefore, given the complexity of the water level time series, pre-processing the water level data first before making the prediction can improve the effectiveness of the prediction.

3.2. Water Level Data Decomposition

Decomposing the original water level time series using DWT first requires determining the appropriate wavelet function and decomposition level. After extensive tests and combining the characteristics of the original water level time series, the Daubechies6 (db6) wavelet function was chosen to decompose the water level data. For the decomposition level L, the empirical equation (Equation. (14)) proposed by the relevant researchers was used to determine .

$$L = int[\,log(\,n_{f(t)})] \qquad (14)$$

where *nf*(*t*) is the series length.

Finally, the db6 wavelet function was used to decompose the original water level three times, and four components were obtained, including one low-frequency sub-series and three high-frequency sub-series, and the results after DWT decomposition are shown in Figure 6. The high-frequency component and low-frequency component reflect the detailed change characteristics and main change trend of the water level time series, respectively. After DWT processing, the original water level is decomposed into several sub-series with different frequency bands, which helps the prediction model to capture the characteristics of the data more easily.

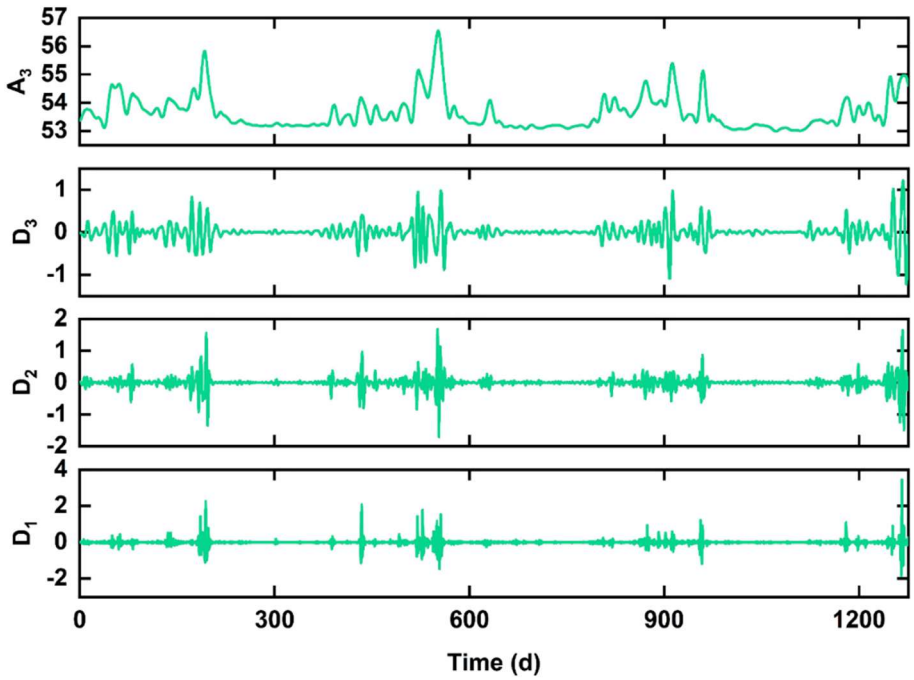

Figure 6. Decomposition results of water level in Le'an River.

3.3. Comparative Analysis with Conventional Single Prediction Models

In this work, BPNN, ELM, and LSTM models are analyzed in comparison with W-LSTM using a consistent dataset. Figure 7 (a,b,c, and f) exhibits the prediction results of individual prediction models. It is obvious from the figures that the constructed W-LSTM model has significant advantages over the prediction results of the three single conventional models, BPNN, ELM, and LSTM. The ELM and LSTM have slightly better prediction results than BPNN because ELM has faster learning speed and better generalization performance compared to BPNN, while the deep learning model LSTM is easier to explore the correlation of the data compared to the shallow learning model BPNN. However, from an overall viewpoint, the prediction accuracy of the above three comparative models is low, especially in the prediction of some peaks and extremes, while W-LSTM still shows high accuracy tracking in the prediction of peaks and extremes.

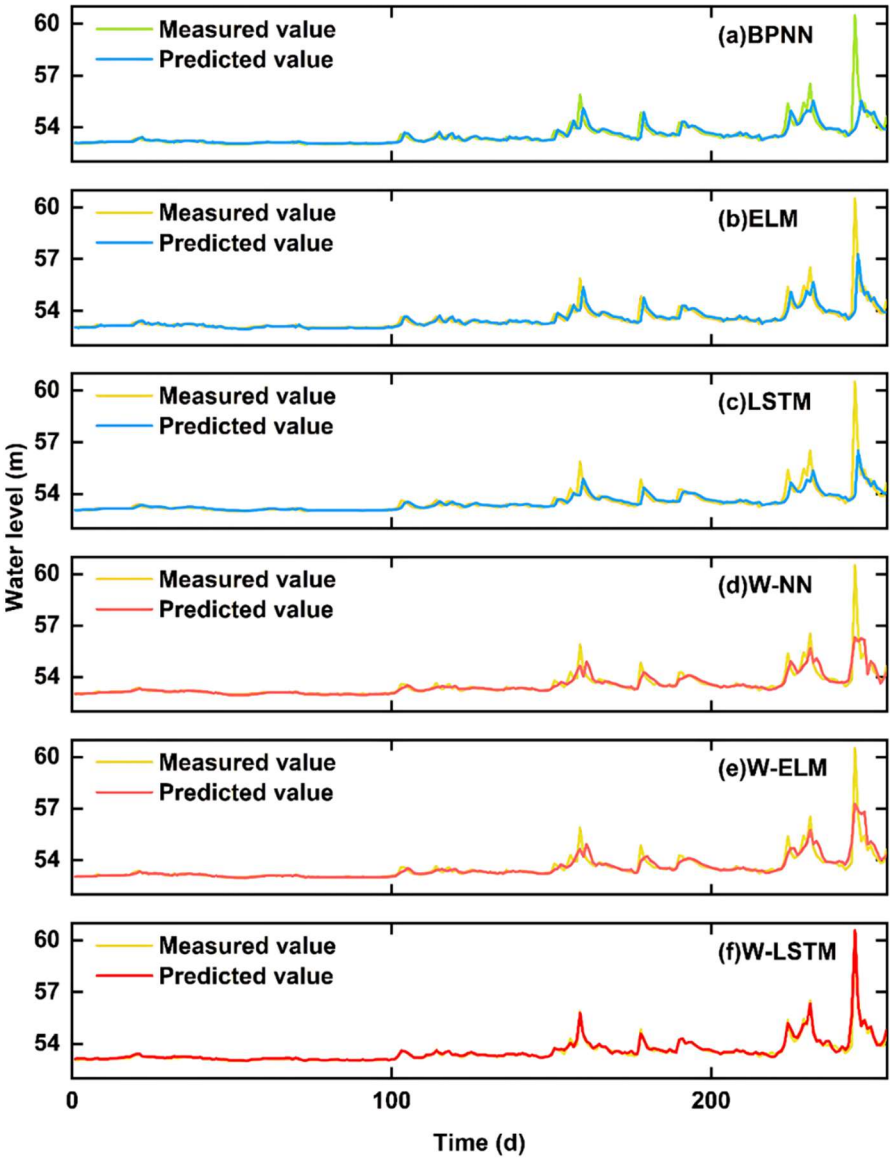

Figure 7. Predicted results for each predictive model.

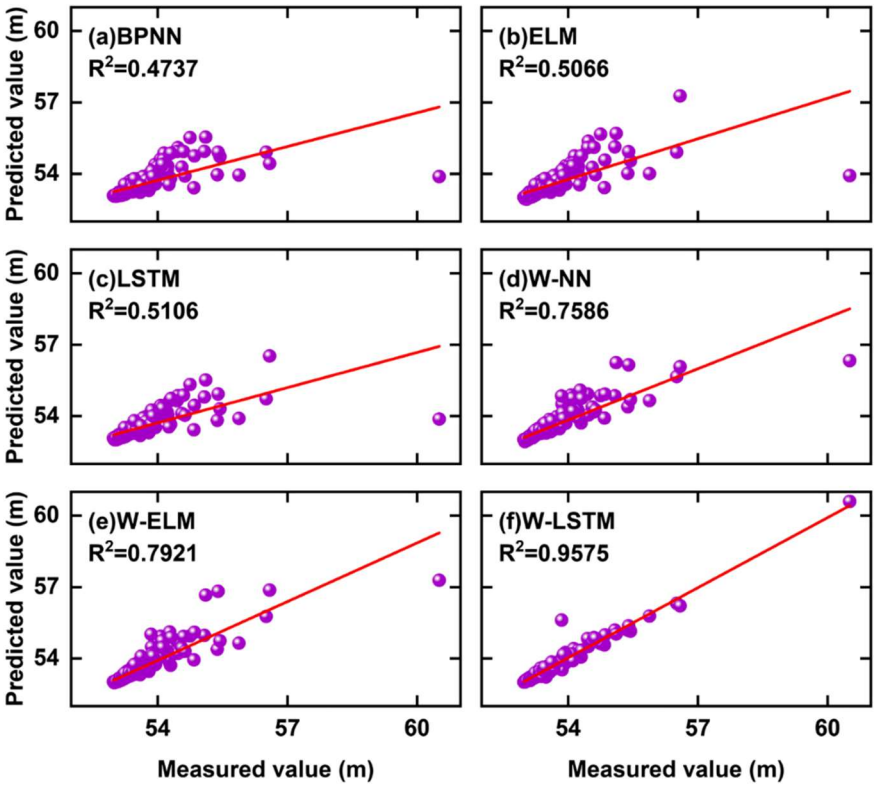

Figure 8. Scatter plots for each predictive model.

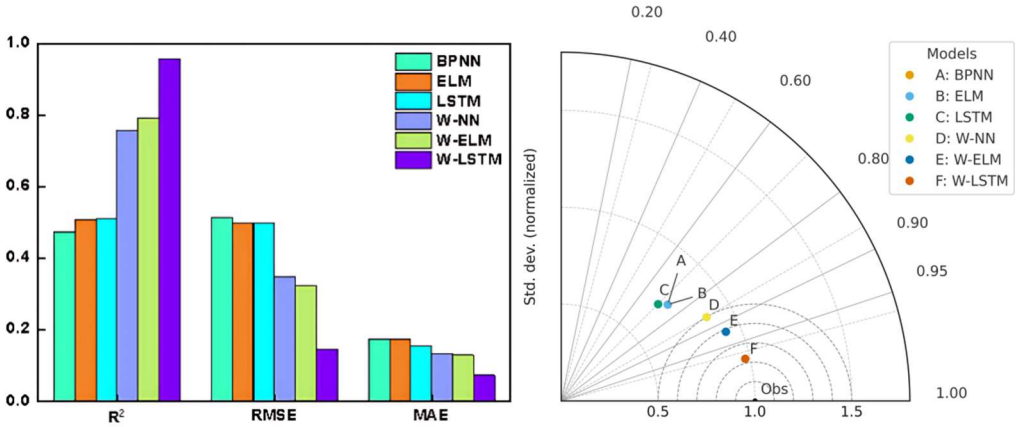

Figure 9. Evaluation criteria for each predictive model.

The scatter plots (Figure 8(a,b,c, and f)) show that the W-LSTM model has a more concentrated scatter distribution of predicted and measured values compared to the other three comparative models, which converge approximately to the diagonal. The slopes of the fitted straight lines of the other three models are far from the diagonal, which also indicates the qualitative perspective that the W-LSTM possesses superior predictive power. According to the evaluation criteria (Figure 9), it can also be seen that BPNN (R^2 = 0.4737, *RMSE* = 0.5153, *MAE* = 0.1745), ELM (R^2 = 0.5066, *RMSE* = 0.4995, *MAE* = 0.1723) and LSTM (R^2 = 0.5106, *RMSE* = 0.4969, *MAE* = 0.1540) are still below acceptable prediction accuracy (R^2 > 0.6). In contrast, the W-LSTM (R^2 = 0.9575, *RMSE* = 0.1465, *MAE* = 0.0737) outperforms the comparative model in all

evaluation criteria, which can also indicate a better prediction performance from the quantitative perspective.

The training and testing values of W-LSTM show minimal deviation, with the test curve closely matching the training curve, whereas the single baseline model exhibits more pronounced divergence. This indicates that the combination of wavelet decomposition (reducing high-frequency noise) and regularization strategies (dropout, early stopping, weight decay, gradient clipping) effectively suppresses overfitting, enabling the model to demonstrate strong generalization capabilities on the holdout set. Furthermore, the relative ranking of models in Figure 7 remains consistent between training and testing, indicating that the improvements achieved by W-LSTM are robust and not attributable to random variations in the training samples.

To quantify the prediction uncertainty for all six models (BPNN, ELM, LSTM, W-ELM, W-LSTM, and W-NN), we calculate PICP (Prediction Interval Coverage Probability) and MPIW (Mean Prediction Interval Width), following the split conformal method. The first 20% of the data is used as the calibration split, and the 95th percentile of absolute residuals is computed as τ. For the evaluation segment, the 95% prediction interval is defined as $[yt-\tau, yt+\tau]$, where yt is the predicted value.

We summarize PICP and MPIW for all models in Table 2, with W-LSTM achieving near-nominal coverage (PICP $\approx$ 0.95) and relatively smaller MPIW, indicating strong validity and compact intervals.

Table 2. Prediction Interval Coverage and Width.

MODEL	PICP	MPIW
BPNN	0.9592	1.3065
ELM	0.9694	1.5602
LSTM	0.9541	1.3868
W-ELM	0.9643	1.5221
W-LSTM	0.949	1.2041
W-NN	0.9592	1.4737

Based on the above analysis, it can be roughly inferred that due to the strong non-stationary and non-linear characteristics in the water level time series, it is difficult for the conventional single model to achieve the expected prediction effect. At the same time, due to the problem of water level data characteristics, the prediction error problem caused by the strong non-linearity existing at its peak and extreme points is also difficult to effectively overcome by a single prediction model.

3.4. Comparative Analysis of Hybrid Prediction Models Coupled with Wavelet Transform

To give further validation of the performance of the constructed method, the hybrid models (W-NN, W-ELM) of BPNN and ELM coupled with discrete wavelet transform are studied in comparison with W-LSTM, using the same dataset as above, and the prediction results are shown in Figure 7(d and e). From the figures, it can be seen that W-NN and W-ELM can roughly track the fluctuations of the water level, but

both still show poor results in tracking the peaks and extreme points. It can also be found that the predicted and measured values of both W-NN and W-ELM are closer to each other overall, especially in the case of small water level fluctuations, but they show large differences locally.

In addition, the scatter plots of both (Figure 8(d and e)) also show that the scatter distribution of the predicted and measured values of W-NN and W-ELM are more concentrated when the water level does not change much in the early period, and the slope of the linear fit approaches the diagonal. However, the distribution of point clusters around the peak and extreme points, where the water level varies greatly, appears to be relatively divergent, which also confirms the limited accuracy of the above analysis in local prediction. Although the few scatters of the W-LSTM also appear to be relatively divergent, this is perfectly acceptable. It is also easy to conclude based on the evaluation criteria (Figure 9) that although the prediction accuracy of W-NN (R^2 = 0.7586, *RMSE* = 0.3489, *MAE* = 0.1321) and W-ELM (R^2 = 0.7921, *RMSE* = 0.3238, *MAE* = 0.1293) meet an acceptable level, with R^2 of both being greater than 0.6, they still achieve not reach the high accuracy of prediction. And compared with W-NN and W-ELM, W-LSTM has improved R^2 by 26.2% and 20.9%, reduced RMSE by 58.0% and 54.8%, and reduced MAE by 44.2% and 43.0%, respectively. This also shows that W-LSTM has a more stable and effective model performance than W-NN and W-ELM, and possesses better applicability among all the compared models. Through the above comparative study, W-NN and W-ELM are effective in predicting water level to some extent, however, the shallow neural network is unable to track the long-term evolution of water level, and there is also a complex relationship between water level and multiple factors, thus making the prediction performance of the model significantly lower.

4. Discussion

4.1. Wavelet Transform-based LSTM Model Enhances Prediction Accuracy

In this study, the wavelet transform and LSTM algorithm were combined to successfully predict the water level changes in the Le'an River. And the comparative analysis confirms that W-LSTM has better model performance.

Considering factors such as weather conditions, hydrological conditions, and the social environment, which lead to the water levels showing complex dynamic changes, greatly increase the difficulty of water level prediction. Considering also the influence of certain local policies also brings many uncertainties to water level prediction. Therefore, for these complex non-linear data sets, it is often difficult to make effective predictions using single methods. In this study, the water level prediction performance of the three different hybrid models clearly improved after pre-processing the original data with the wavelet transform. This is because wavelet transform can simultaneously extract frequency and time information present in water level time series at high resolution in both the frequency and time domains. After wavelet transform processing, the noise in the data set is effectively removed, which can make it easier for the prediction model to extract internally useful information to capture and learn its

dynamic features, thus effectively improving the predictive power and applicability of the model.

In addition, LSTM has the best prediction performance among all wavelet transform-based models (BPNN, ELM, and LSTM). Because of its internal structure, LSTM can retain effective information for a long time while discarding useless information, thus ensuring that the flow of information will not be easily changed. Therefore, LSTM is easier to extract the features of the decomposed water level series. At the same time, LSTM is able to infer new water level features from the finite set of features contained in the training set used, and this advantage of deep learning is also the difficulty faced by BPNN and ELM. Therefore, W-LSTM has a more powerful prediction ability in this water level prediction work.

4.2. Limitations and Prospects

While this study has achieved the expected results, there are still some limitations that need to be further explored. First of all, the area of this study is relatively single, and the water level changes caused by geographical location and environmental factors vary greatly in different regions, so several representative areas can be selected for comprehensive studies in the future to confirm the practicality of the constructed method and reach more thorough conclusions. In addition, this study did not achieve perfect prediction accuracy because the high-frequency component of the wavelet transform decomposition still has some noisy information, so further optimization is needed for the subsequent work, such as the consideration of the secondary decomposition method. Finally, this study relies on water level data characteristics from the perspective of water level prediction, so it has limitations in application. Future work can be devoted to the further improvement of the method in order to apply it to a wider range of scenarios, which can also provide new perspectives for prediction work.

5. Conclusions

To address the challenge of water level prediction for effective water environment management, this study proposes the W-LSTM hybrid model, which combines wavelet transform with the deep learning model LSTM. This approach enhances the accuracy and stability of water level predictions compared to traditional single models. By utilizing wavelet transform, which effectively decomposes the water level time series into components with distinct features, and combining it with the deep learning capabilities of LSTM, the proposed model demonstrates superior performance in capturing long-term dependencies in the data. The empirical results confirm that W-LSTM outperforms conventional models, such as BPNN, ELM, and LSTM, in terms of prediction accuracy and generalization ability. The model's ability to accurately predict water levels, especially during extreme events, highlights its robustness and potential for real-world applications in water resource management and flood prevention.

In conclusion, the W-LSTM model represents a powerful tool for water level prediction, providing valuable insights for flood control and waterway management. Future research could focus on expanding this model to other geographical regions

with different hydrological characteristics and further optimizing its performance using more advanced techniques, such as multi-step ahead prediction and k-fold cross-validation.

Author contributions: B.L.: Methodology, Formal analysis, Software, Writing—Original Draft. J.Y.: Writing—Review and Editing, Resources. J.Z.: Review and Editing, Supervision, Resources. C.L.: Resources. Y.Z.: Supervision. Y.L.: Review and Editing. L.G.: Review and Editing. L.L.: Review and Editing. R.L.: Review and Editing. W.Y.: Writing—Review and Editing, Resources, Conceptualization, Investigation.

Funding: This work was jointly supported by the National Natural Science Foundation of China (42077156, 52121006, and 22006041), Chongqing Basic and Frontier Research Project (cstc2020jcyj-msxmX0176), Guangdong Basic and Applied Basic Research Foundation (2020A1515111128, 202002030169, and 2020A1515011130).

Acknowledgments: Here, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflict of interest: The authors declare no conflict of interest.

Reference

1. Luo P., Luo M., Li F., Qi X., Huo A., Wang Z., He B., Takara K., Nover D., Wang Y. Urban flood numerical simulation: Research, methods and future perspectives. Environmental Modelling & Software. 2022, 156, 105478. https://doi.org/10.1016/j.envsoft.2022.105478
2. Kim D., Lee J., Kim J., Lee M., Wang W., Kim H.S. Comparative analysis of long short-term memory and storage function model for flood water level forecasting of Bokha stream in NamHan River, Korea. Journal of Hydrology. 2022, 606, 127415. https://doi.org/10.1016/j.jhydrol.2021.127415
3. Wang X., Xia J., Zhou M., Deng S., Li Q. Assessment of the joint impact of rainfall and river water level on urban flooding in Wuhan City, China. Journal of Hydrology. 2022, 613, 128419. https://doi.org/10.1016/j.jhydrol.2022.128419
4. Hinge G., Hamouda M.A., Long D., Mohamed M.M. Hydrologic utility of satellite precipitation products in flood prediction: A meta-data analysis and lessons learnt. Journal of Hydrology. 2022, 612, 128103. https://doi.org/10.1016/j.jhydrol.2022.128103
5. Hu L., Zhang Q., Wang G., Singh V.P., Wu W., Fan K., Shen Z. Flood disaster risk and socioeconomy in the Yellow River Basin, China. Journal of Hydrology: Regional Studies. 2022, 44, 101272. https://doi.org/10.1016/j.ejrh.2022.101272
6. Guo X., Cheng J., Yin C., Li Q., Chen R., Fang J. The extraordinary Zhengzhou flood of 7/20, 2021: How extreme weather and human response compounding to the disaster. Cities. 2023, 134, 104168. https://doi.org/10.1016/j.cities.2022.104168
7. Rigos A., Krommyda M., Tsertou A., Amditis A. A polynomial neural network for river's water-level prediction. SN Applied Sciences. 2020, 2, 529. https://doi.org/10.1007/s42452-020-2328-9
8. Yuan Z., Liu J., Liu Y., Zhang Q., Li Y., Li Z. A two-stage modelling method for multi-station daily water level prediction, Environmental Modelling & Software. 2022, 156, 105468. https://doi.org/10.1016/j.envsoft.2022.105468
9. Cho M., Kim C., Jung K., Jung H. Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)–Gated Recurrent Unit (GRU) Method for Flood Prediction. Water. 2022, 14, 2221. https://doi.org/10.3390/w14142221
10. Pan M., Zhou H., Cao J., Liu Y., Hao J., Li S., Chen C.-H. Water level prediction model based on GRU and CNN. IEEE Access. 2020, 8, 60090–60100. https://doi.org/10.1109/ACCESS.2020.2982433

11. Liu Z., Cheng L., Lin K., Cai H. A hybrid bayesian vine model for water level prediction. Environmental Modelling & Software. 2021, 142, 105075. https://doi.org/10.1016/j.envsoft.2021.105075
12. Wang Q., Wang S. Machine learning-based water level prediction in Lake Erie. Water. 2020, 12, 2654. https://doi.org/10.3390/w12102654
13. Yang D., Herath S., Musiake K. A hillslope-based hydrological model using catchment area and width functions. Hydrological Sciences Journal. 2002, 47, 49–65. https://doi.org/10.1080/02626660209492907
14. Cao Y.H., Ye X.F., Lv G.H. Study on Water Level Prediction in the Ganjiang Valley. In: Applied Mechanics and Materials. Trans Tech Publications: Baech, Switzerland, 2012. Vol. 212, pp. 417–422.
15. El-Diasty M., Al-Harbi S. Development of wavelet network model for accurate water levels prediction with meteorological effects. Applied Ocean Research. 2015, 53, 228–235. https://doi.org/10.1016/j.apor.2015.09.008
16. Connelly B.A., Braatz D.T., Halquist J.B., DeWeese M.M., Larson L., Ingram J.J. Advanced hydrologic prediction system. Journal of Geophysical Research: Atmospheres. 1999, 104, 19655–19660. https://doi.org/10.1029/1999jd900051
17. Xu S., Huang W. Estimating extreme water levels with long-term data by GEV distribution at Wusong station near Shanghai city in Yangtze Estuary. Ocean Engineering. 2011, 38, 468–478. https://doi.org/10.1016/j.oceaneng.2010.11.022
18. Xing W.Y., Bai Y.L., Ding L., Yu Q.H., Song W. Application of a hybrid model based on GA–ELMAN neural networks and VMD double processing in water level prediction. Journal of Hydroinformatics. 2022, 24, 818–837. https://doi.org/10.2166/hydro.2022.016
19. Truong V.-H., Ly Q.V., Le V.-C., Vu T.-B., Tran T.-T., Goethals P. Machine learning-based method for forecasting water levels in irrigation and drainage systems. Environmental Technology & Innovation. 2021, 23, 101762. https://doi.org/10.1016/j.eti.2021.101762
20. Yu P.-S., Chen S.-T., Chang I.-F. Support vector regression for real-time flood stage forecasting. Journal of hydrology. 2006, 328, 704–716. https://doi.org/10.1016/j.jhydrol.2006.01.021
21. Nordin N.F.C., Mohd N.S., Koting S., Ismail Z., Sherif M., El-Shafie A. Groundwater quality forecasting modelling using artificial intelligence: A review. Groundwater for Sustainable Development. 2021, 14, 100643. https://doi.org/10.1016/j.gsd.2021.100643
22. Soleymani S.A., Goudarzi S., Anisi M.H., Hassan W.H., Idris M.Y.I., Shamshirband S., Noor N.M., Ahmedy I. A novel method to water level prediction using RBF and FFA. Water Resources Management. 2016, 30, 3265–3283. https://doi.org/10.1007/s11269-016-1347-1
23. Deo R.C., Şahin M. An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland. Environmental monitoring and assessment. 2016, 188, 90. https://doi.org/10.1007/s10661-016-5094-9
24. Chen N., Xiong C., Du W., Wang C., Lin X., Chen Z. An improved genetic algorithm coupling a back-propagation neural network model (IGA-BPNN) for water-level predictions. Water. 2019, 11, 1795. https://doi.org/10.3390/w11091795
25. Ren T., Liu X., Niu J., Lei X., Zhang Z. Real-time water level prediction of cascaded channels based on multilayer perception and recurrent neural network. Journal of Hydrology. 2020, 585, 124783. https://doi.org/10.1016/j.jhydrol.2020.124783
26. Yang C.-H., Wu C.-H., Hsieh C.-M. Long short-term memory recurrent neural network for tidal level forecasting. IEEE Access. 2020, 8, 159389–159401. https://doi.org/10.1109/ACCESS.2020.3017089
27. Liu Y., Wang H., Feng W., Huang H. Short term real-time rolling forecast of urban river water levels based on lstm: a case study in fuzhou city, China. International Journal of Environmental Research and Public Health. 2021, 18, 9287. https://doi.org/10.3390/ijerph18179287
28. Li P., Hua P., Gui D., Niu J., Pei P., Zhang J., Krebs P. A comparative analysis of artificial neural networks and wavelet hybrid approaches to long-term toxic heavy metal prediction. Scientific reports. 2020, 10, 13439. https://doi.org/10.1038/s41598-020-70438-8
29. Sang Y.-F., Wang D., Wu J.-C., Zhu Q.-P., Wang L. The relation between periods' identification and noises in hydrologic series data. Journal of Hydrology. 2009, 368, 165–177. https://doi.org/10.1016/j.jhydrol.2009.01.042
30. Joo T.W., Kim S.B. Time series forecasting based on wavelet filtering. Expert Systems with Applications. 2015, 42, 3868–3874. https://doi.org/10.1016/j.eswa.2015.01.026
31. Rhif M., Ben Abbes A., Farah I.R., Martínez B., Sang Y. Wavelet transform application for/in non-stationary time-series analysis: a review. Applied Sciences. 2019, 9, 1345. https://doi.org/10.3390/app9071345

32. Wu C., Zhang X., Wang W., Lu C., Zhang Y., Qin W., Tick G.R., Liu B., Shu L. Groundwater level modeling framework by combining the wavelet transform with a long short-term memory data-driven model. Science of The Total Environment. 2021, 783, 146948. https://doi.org/10.1016/j.scitotenv.2021.146948
33. Li Z., Sun Z., Liu J., Dong H., Xiong W., Sun L., Zhou H. Prediction of river sediment transport based on wavelet transform and neural network model. Applied Sciences. 2022, 12, 647. https://doi.org/10.3390/app12020647
34. Belyakova P., Moreido V., Tsyplenkov A., Amerbaev A., Grechishnikova D., Kurochkina L., Filippov V., Makeev M. Forecasting Water Levels in Krasnodar Krai Rivers with the Use of Machine Learning. Water Resources. 2022, 49, 10–22. https://doi.org/10.1134/S0097807822010043
35. Altunkaynak A., Kartal E. Performance comparison of continuous wavelet-fuzzy and discrete wavelet-fuzzy models for water level predictions at northern and southern boundary of Bosphorus. Ocean Engineering. 2019, 186, 106097. https://doi.org/10.1016/j.oceaneng.2019.06.002
36. Liu Y., Guan L., Hou C., Han H., Liu Z., Sun Y., Zheng M. Wind power short-term prediction based on LSTM and discrete wavelet transform. Applied Sciences. 2019, 9, 1108. https://doi.org/10.3390/app9061108
37. Nanda T., Sahoo B., Beria H., Chatterjee C. A wavelet-based non-linear autoregressive with exogenous inputs (WNARX) dynamic neural network model for real-time flood forecasting using satellite-based rainfall products. Journal of Hydrology. 2016, 539, 57–73. https://doi.org/10.1016/j.jhydrol.2016.05.014
38. Hochreiter S., Schmidhuber J. Long short-term memory. Neural computation. 1997, 9, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
39. Zhang Q., Wang R., Qi Y., Wen F. A watershed water quality prediction model based on attention mechanism and Bi-LSTM. Environmental Science and Pollution Research. 2022, 29, 75664–75680. https://doi.org/10.1007/s11356-022-21115-y
40. Liu P., Wang J., Sangaiah A.K., Xie Y., Yin X. Analysis and prediction of water quality using LSTM deep neural networks in IoT environment. Sustainability. 2019, 11, 2058. https://doi.org/10.3390/su11072058
41. 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. Journal of Hydrology. 2021, 598, 126196. https://doi.org/10.1016/j.jhydrol.2021.126196