Proceedings of the 3rd International Conference on Functional Materials and Civil Engineering
DOI: 10.54254/2755-2721/2025.29610

Urban Waterlogging Volume Prediction Based on Differential Evolution Algorithm and Bidirectional Long Short-Term Memory Network Optimization Transformer Algorithm

Changtao Deng[1*], Xiaowei Li[2], Dan Feng[3]

[1]Pearl River Water Resources Research Institute, Guangzhou, China
[2]Hainan Provincial Institute of Modern Emergency Management, Haikou, China
[3]Hainan Provincial Center for Disease Control and Prevention, Haikou, China
**Corresponding Author. Email: danchun2020@126.com*

Abstract. Urban waterlogging is a typical complex disaster in the process of modern urbanization, triggered by the superimposition of multiple structural contradictions. To enhance the accuracy and timeliness of water accumulation depth prediction and provide core technical support for the full-chain prevention and control of urban waterlogging monitoring - early warning - dispatching, this paper proposes the DE-Transformer-BiLSTM regression algorithm. The research first conducted a correlation analysis and then tested and verified it through multiple machine learning algorithms. The results showed that the algorithm performed the best on all core regression indicators: Compared with traditional integrated models such as ExtraTrees and Random Forest, its MSE is reduced by 43.7%, RMSE is reduced by 25.0%, MAE is reduced by 23.4%, MAPE is reduced by 11.9%, and R^2 is increased to 0.988, significantly optimizing error control and data interpretation capabilities. Compared with other deep learning models such as SAIS-CNN-BILSTM and NRBO-Transformer-BiLSTM, with better hyperparameter optimization effects and stronger long-range dependency and bidirectional temporal feature modeling capabilities, it leads by 37.4% to 51.9% in indicators such as MSE and RMSE, and maintains the lowest MAPE. Compared with the second-tier models such as Decision Tree and CatBoost, and the third-tier models such as KNN, the advantages are more significant. The MSE is only 26.6% of CatBoost and 1.5% of KNN, and the R^2 is close to the theoretical maximum value of 1. The advantage of this algorithm stems from its unique architecture. The DE algorithm efficiently optimizes hyperparameters to reduce labor costs, the Transformer captures long-range features, and BiLSTM extracts bidirectional temporal information. The combination of the three not only solves the problem of difficult adjustment of hyperparameters in deep learning models but also makes up for the deficiency of complex feature modeling in traditional models. It provides a reliable technical solution for the precise prevention and control of urban flooding.

Keywords: Urban flooding, Transformer, BiLSTM

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.29610

1. Introduction

Urban waterlogging has become a prominent compound disaster in the process of modern urbanization, and its formation stems from the superposition of multiple structural contradictions. With the expansion of urban construction land, the hardened surface has replaced natural soil, resulting in a decline of over 60% in rainwater infiltration capacity. Moreover, more than 60% of urban drainage systems in China are designed to meet standards lower than the annual standard, and in some old urban areas, the pipe networks even follow the standards from the early days of the founding of the People's Republic of China [1]. Under the backdrop of global climate change, the average number of days with heavy rainfall across the country has increased by 12% in the past decade compared to the end of the last century. Extreme rainfall events have occurred frequently. Coupled with problems such as the encroachment of river and lake water system regulation and storage space, the risk of urban flooding has continued to rise. The traditional monitoring system has obvious shortcomings. Its reliance on scattered single-point equipment leads to insufficient coverage of the entire area [2].

Machine learning algorithms have provided a breakthrough solution for urban waterlogging monitoring, promoting the transformation of monitoring models from passive perception to active early warning [3]. Compared with traditional hydrological models, machine learning has a powerful ability to fuse multi-source data and perform nonlinear mapping, and can integrate more than 21 types of influencing factors such as meteorology, hydrology, topography and drainage systems. Among mainstream algorithms, Random Forest and LightGBM are adept at risk level classification and key influencing factor identification, while deep learning models like LSTM can capture the temporal correlation between rainfall and water accumulation, enabling dynamic prediction of water accumulation depth [4]..

To break through the above bottlenecks, this paper proposes the DE-Transformer-BiLSTM regression algorithm. This algorithm specifically integrates the advantages of multiple models: BiLSTM, with its bidirectional time series modeling capability, accurately captures the short-term local correlations of factors such as rainfall intensity and duration, compensating for the memory deficiency of traditional RNNS. The Transformer module effectively extracts the global dependency relationships of spatial features such as terrain and drainage systems through its self-attention mechanism, solving the problem of weak global perception in BiLSTM. The differential evolution algorithm (DE) is introduced to globally optimize the hyperparameters of the model.

2. Data sources

This dataset collects urban waterlogging monitoring data from January 2020 to December 2023. Data collection covers 98 flood-prone areas. By combining intelligent monitoring terminals with manual inspection records, the accuracy of the data is ensured. The basic meteorological data comes from the network of 18 automatic weather stations of the city meteorological bureau. The operation data of the drainage system is synchronized in real time from the municipal smart water management platform. Geographic information data is based on the 1:5000 urban GIS system. The soil data is from the soil monitoring network of the Municipal Natural Resources Bureau. The depth of waterlogging is verified through both laser level gauges and manual measurement.

DOI: 10.54254/2755-2721/2025.29610

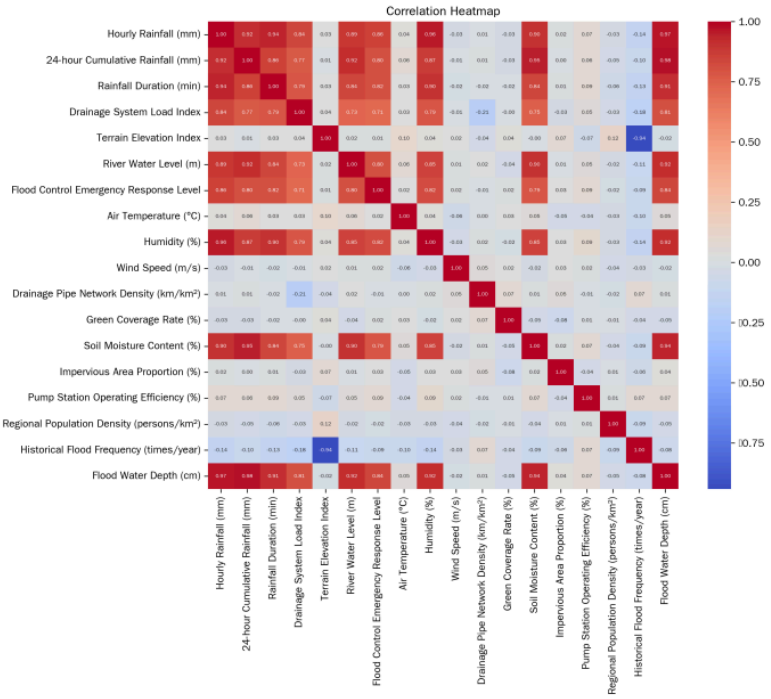


Figure 1. Correlation heat map

In the preprocessing stage, the time series interpolation method was adopted to fill in the missing values, the IQR method was used to eliminate outliers, and the variables of different magnitudes were standardized. Eventually, 572 valid records were screened out, covering different seasons, weather types and urbanized areas.First, conduct a correlation analysis on each variable and draw a correlation heat map, as shown in Figure 1. From these correlation coefficients, it can be inferred that variables related to rainfall such as hourly rainfall (mm), cumulative rainfall in the previous 24 hours (mm), and rainfall duration (min), as well as variables such as river water level (m), humidity (%), and soil moisture content (%), have a strong positive correlation with the depth of waterlogging (cm). It indicates that these factors have a relatively significant positive impact on the depth of urban waterlogging. That is, when the values of these factors increase, the depth of urban waterlogging may deepen accordingly.

3. Method

3.1. Differential Evolution algorithm

Differential Evolution (DE) algorithm is a population-based random search optimization algorithm belonging to the evolutionary algorithm family. Its core advantage lies in its ability to efficiently solve complex optimization problems in continuous Spaces without relying on gradient information [5]. Its essence is to guide the population to gradually approach the optimal solution by simulating the mutation-cross-selection process in biological evolution, and it is particularly suitable for high-dimensional, nonlinear or multimodal optimization tasks [6].

The core process of DE is carried out around four major steps: First, population initialization, randomly generating an initial population of size NP, with each individual corresponding to a solution vector of the optimization problem, and the vector dimension is consistent with the number of optimization variables; The second is the mutation operation, which is a signature step of DE -

DOI: 10.54254/2755-2721/2025.29610

generating a mutation vector based on the difference information of any three different individuals in the population [7].

3.2. BiLSTM

BiLSTM is an extended structure of LSTM, with its core being to address the limitation of traditional LSTM that it can only process time series data in one direction. Traditional LSTM alleviates the vanishing gradient problem of RNN through gating mechanisms and can capture the long-term dependencies of time series data. However, it can only pass information from the start to the end of the sequence and cannot utilize the context of future moments. In many scenarios, the current state needs to rely on both past and future information simultaneously. BiLSTM is precisely designed to meet this demand [8]. The network structure of BiLSTM is shown in Figure 2.

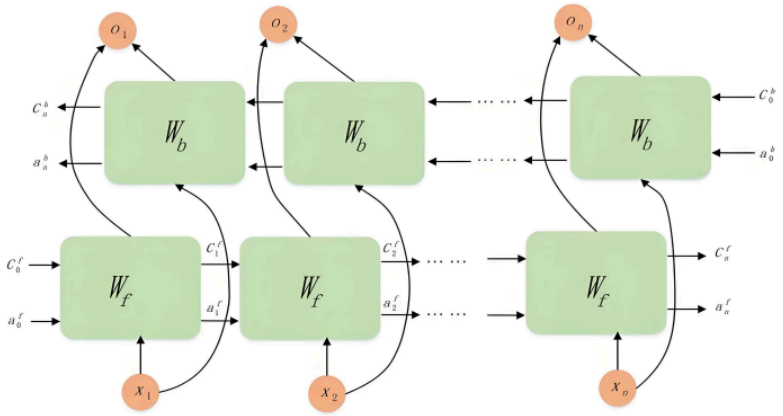


Figure 2. The network structure of BiLSTM

The core structure of BiLSTM consists of two independent LSTM modules: Forward LSTM and Backward LSTM. When dealing with time series, the forward LSTM starts from the first time step of the sequence and processes backward one by one to capture the time series correlation from the past to the present. Backward LSTM starts from the last time step of the sequence and processes backward to capture the temporal correlation from the future to the present. The two modules operate in parallel, concatenating and fusing their respective output feature vectors at each time step to form the final feature containing bidirectional temporal information, providing a more comprehensive temporal context for subsequent tasks. This bidirectional structure gives BiLSTM a unique advantage in timing-related tasks: it can retain LSTM's ability to handle long-term dependencies while avoiding context loss caused by one-way information transmission [8].

3.3. Transformer

Transformer is a deep learning model proposed by the Google team [9]. Its core breakthrough lies in its complete construction based on the attention mechanism, which breaks away from the temporal dependence limitations of models such as recurrent neural networks (RNN) and LSTM, achieving efficient parallel processing of temporal data and more accurately capturing long-distance dependencies [10]. The network structure of the Transformer is shown in Figure 3.

DOI: 10.54254/2755-2721/2025.29610

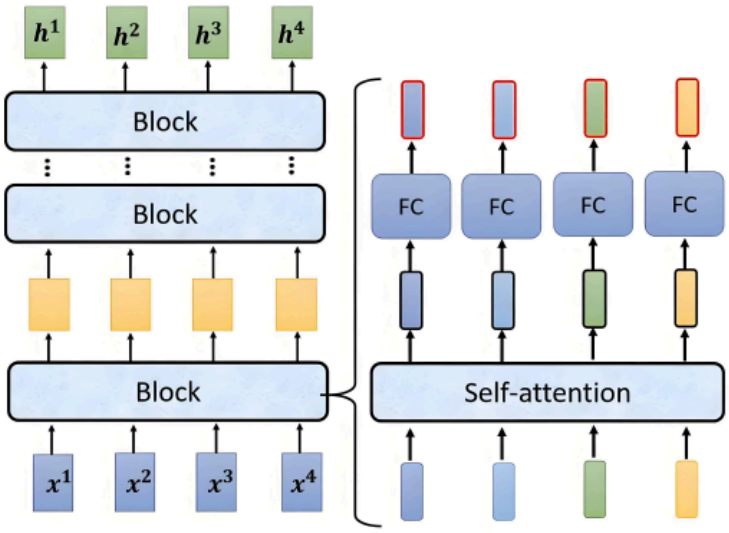


Figure 3. The network structure of the Transformer

3.4. DE-Transformer-BiLSTM

The DE-Transformer-BILSTM regression algorithm is a fusion model designed for the regression task of predicting the depth of urban waterlogging accumulation. The core lies in the organic synergy of the differential evolution algorithm (DE), Transformer, and bidirectional Long Short-Term memory Network (BiLSTM). Enhance the ability to model temporal and spatial features and the stability of model performance.

Firstly, by leveraging Transforme's multi-head self-attention mechanism, global dependency capture is conducted on the spatial correlation features related to urban waterlogging. The correlation weights between each feature are calculated through Query, Key, and Value vectors, focusing on key influencing factors to generate feature vectors containing global spatial information. Secondly, the feature vector is input into the BiLSTM module. With the help of its forward LSTM, the historical cumulative influence of time series data such as rainfall intensity and soil moisture content is captured, and the backward LSTM captures the future short-term trend. After bidirectional fusion, the temporal-spatial joint features are generated. Finally, the DE algorithm is introduced to globally optimize the model's hyperparameters. Through an iterative process of mutation, crossover, and selection, the optimal parameter configuration is screened out to avoid the local optimum trap of manual parameter adjustment.

4. Result

In this paper, the DE-Transformer-BiLSTM regression algorithm is used to predict urban waterlogging. In terms of parameter Settings, the number of search agents is 8, the maximum number of iterations is 6, the optimization parameter dimension is 3, and the lower bounds of the parameters are respectively the learning rate 1e-3, the hidden layer node 10, and the regularization coefficient 1e-4. The upper bounds are respectively the learning rate 1e-2, the hidden layer node 30, and the regularization coefficient 1e-1; The model structure parameters include the number of heads in the Transformer part being 4, the maximum position encoding being 512, the number of key channels for each head being 128, the number of hidden layer nodes in BiLSTM being optimized,

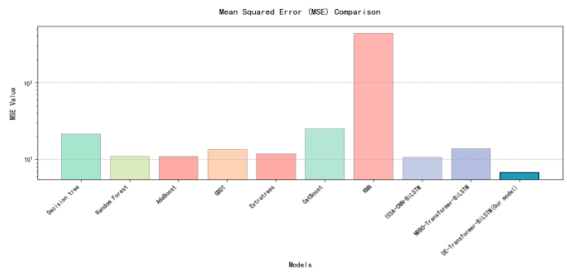
DOI: 10.54254/2755-2721/2025.29610

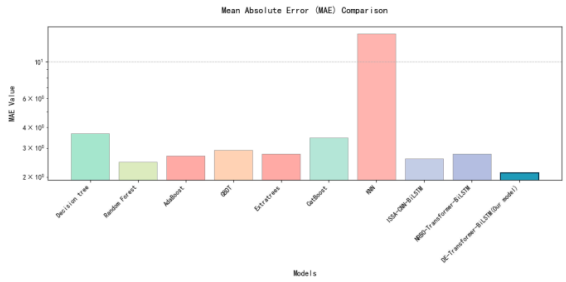
and the dropout rate being 0.01. The training parameters include a maximum of 1000 training rounds, a batch size of 256, a learning rate decline factor of 0.1, a learning rate decline period of 850, and a gradient clipping threshold of 10. This paper first conducted a comparative experiment The comparison models selected Decision tree, Random Forest, AdaBoost, GBDT, Extratrees, CatBoost, KNN, ISSA-CNN-BiLSTM and NRBO-Transformer-BiLSTM The experimental results are shown in Table 1.

Table 1. The experimental results

Model	MSE	RMSE	MAE	MAPE	R^2
Decision tree	21.598	4.647	3.674	18.109	0.969
Random Forest	10.951	3.309	2.475	12.68	0.981
AdaBoost	10.938	3.307	2.697	13.621	0.981
GBDT	13.59	3.686	2.909	14.622	0.979
Extratrees	11.925	3.453	2.768	13.769	0.983
CatBoost	25.175	5.018	3.465	14.205	0.953
KNN	436.313	20.888	14.864	60.146	0.174
ISSA-CNN-BiLSTM	10.715	3.273	2.583	17.051	0.98
NRBO-Transformer-BiLSTM	13.958	3.736	2.766	14.076	0.975
DE-Transformer-BiLSTM(Our model)	6.705	2.589	2.12	12.135	0.988

A bar chart comparison and analysis of each indicator was conducted, as shown in Figure 4. DE-Transformer-BiLSTM performed the best on all core regression metrics. Compared with traditional ensemble models such as ExtraTrees and Random Forest, its MSE decreased by 43.7%, RMSE decreased by 25.0%, MAE decreased by 23.4%, and MAPE decreased by 11.9%. The R^2 has been raised to 0.988, significantly optimizing the error control and data interpretation capabilities. Compared with other deep learning models such as SAIS-CNN-BILSTM and NRBO-Transformer-BiLSTM, it leads by 37.4% to 51.9% respectively in metrics such as MSE and RMSE, thanks to its superior hyperparameter optimization effect and stronger modeling ability for long-range dependencies and bidirectional temporal features. MAPE also remained at the lowest level. For second-tier models such as Decision Tree and CatBoost, and third-tier models like KNN, the advantages are even more significant. The MSE is only 26.6% of CatBoost and 1.5% of KNN, and the R^2 is close to the theoretical maximum value of 1, far exceeding other low-performance models. This leading position stems from its unique architecture The bar chart comparing the indicators of each model is shown in Figure 4.





DOI: 10.54254/2755-2721/2025.29610

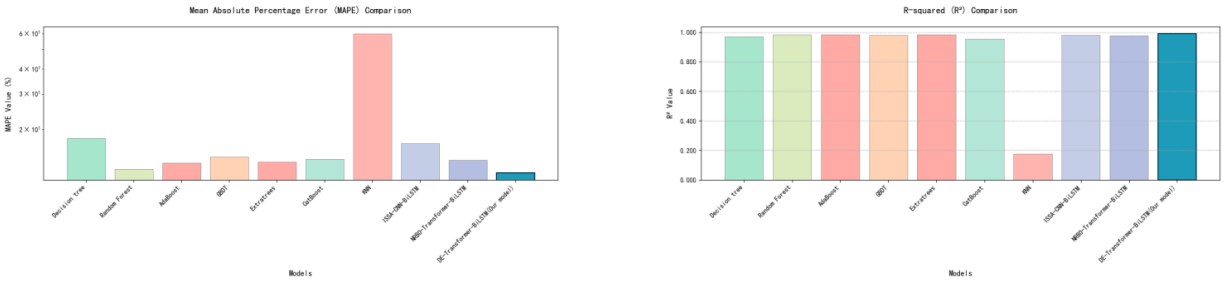


Figure 4. The bar chart comparing the indicators of each model

5. Conclusion

In the process of modern urbanization, urban waterlogging, as a typical complex disaster, has become increasingly prominent. Its occurrence is the result of the combined effect of multiple structural contradictions. The DE-Transformer-BiLSTM regression algorithm developed in this paper can effectively improve the accuracy and timeliness of water accumulation depth prediction, providing key technical support for the full-chain prevention and control of urban waterlogging monitoring, early warning and dispatching. The study first conducted a correlation analysis and then tested through multiple machine learning algorithms. It was found that this algorithm performed best on all core regression indicators. Compared with traditional ensemble models such as ExtraTrees and Random Forest, Its MSE decreased by 43.7%, RMSE by 25.0%, MAE by 23.4%, MAPE by 11.9%, and R^2 increased to 0.988, significantly optimizing the error control and data interpretation capabilities. Compared with other deep learning models such as SAIS-CNN-BILSTM and NRBO-Transformer-BiLSTM, it leads by 37.4% to 51.9% respectively in indicators such as MSE and RMSE, thanks to its superior hyperparameter optimization effect and stronger long-range dependency and bidirectional temporal feature modeling capabilities. MAPE also remained at the lowest level. Compared with second-tier models such as Decision Tree and CatBoost, and third-tier models such as KNN, the advantages are more obvious. The MSE is only 26.6% of CatBoost and 1.5% of KNN, and the R^2 is close to the theoretical maximum value of 1, far exceeding other low-performance models. This advantage is attributed to its unique architecture. The DE algorithm can efficiently optimize hyperparameters to reduce labor costs, the Transformer can capture long-range features, and BiLSTM can extract bidirectional temporal information. The combination of the three not only solves the problem of difficult adjustment of hyperparameters in deep learning models but also makes up for the deficiency of traditional models in complex feature modeling. It provides more reliable technical means for the prevention and control of urban waterlogging, helping to enhance the city's capacity to deal with waterlogging disasters.

References

[1] Sumathy, V., and G. Navamani. "AI-Driven Personalized Learning: Enhancing Student Success through Adaptive Technologies." Library of Progress-Library Science, Information Technology & Computer 44.3 (2024).
[2] Yekollu, Roop Kumar, et al. "AI-driven personalized learning paths: Enhancing education through adaptive systems." International Conference on Smart data intelligence. Singapore: Springer Nature Singapore, 2024.
[3] Tariq, Muhammad Usman. "AI-Driven Personalized Learning-Enhancing Student Engagement and Outcomes." AI and Learning Analytics in Distance Learning. IGI Global Scientific Publishing, 2025. 151-178.
[4] Yu, Haixia, et al. "Artificial intelligence-based quality management and detection system for personalized learning." Journal of Interconnection Networks 22.Supp02 (2022): 2143004.

DOI: 10.54254/2755-2721/2025.29610

[5] Maghsudi, Setareh, et al. "Personalized education in the artificial intelligence era: what to expect next." IEEE Signal Processing Magazine 38.3 (2021): 37-50.
[6] Castro, Gina Paola Barrera, et al. "Harnessing AI for Education 4.0: Drivers of Personalized Learning." Electronic Journal of e-Learning 22.5 (2024): 1-14.
[7] Tavakoli, Mohammadreza. "Hybrid human-AI driven open personalized education." (2023).
[8] Vashishth, Tarun Kumar, et al. "AI-driven learning analytics for personalized feedback and assessment in higher education." Using traditional design methods to enhance AI-driven decision making. IGI Global Scientific Publishing, 2024. 206-230.
[9] Lannelongue, Loïc, and Michael Inouye. "Pitfalls of machine learning models for protein–protein interaction networks." Bioinformatics 40.2 (2024): btae012.
[10] Gm, Dhananjaya, et al. "A digital recommendation system for personalized learning to enhance online education: A review." IEEE Access 12 (2024): 34019-34041.