Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 12, No. 3, 2025

Study on Urban Flood Risk Assessment and Precipitation Prediction Based on SA-Informer Model

Huipeng Xie [1], Zhengbin Xu [1, *], Jiajun Lv [1], Yujia Dai [2]

[1] School of Mathematical Sciences, Chengdu University of Technology, Chengdu, Sichuan, China
[2] College of Materials and Chemistry & Chemical Engineering, Chengdu University of Technology, Chengdu, Sichuan, China
*** Corresponding author:** Zhengbin Xu (Email: XZB_CDUT@163.com)

Abstract: This study focuses on urban flood risk assessment and precipitation prediction, constructs a technical framework centred on the SA-Informer model, and systematically elaborates the algorithmic innovation and application value of this model in precipitation prediction. Firstly, at the risk assessment level, this study integrates multi-source data indicators based on the Pressure-State-Response (P-S-R) model, and confirms that heavy precipitation, as a core driving factor, is significantly associated with the risk level jumping through the AHP-CRITIC combined weighting algorithm, the improved ternary coupling coordination degree model, and the stepwise regression analysis method, which identifies the most influential indicators affecting system coordination. This lays the foundation of causal logic for the construction of the prediction model. Further, the SA-Informer model is innovatively proposed, which combines the global search capability of the simulated annealing algorithm and the long time series processing advantage of the Informer model, and by dynamically adjusting the learning rate, input sequence length and other parameters, the nonlinear characteristics and seasonal patterns of precipitation data are effectively captured. Experiments show that the model significantly improves the prediction accuracy compared with the traditional time series prediction model, and provides highly reliable feed-forward information for flood warning. Based on data fusion and model algorithms, the framework provides a complete solution for urban flood prevention and control, and its methodology and model architecture are highly adaptable and scalable, which can be extended to other climate-sensitive cities in the fields of disaster prediction, risk management and sustainable development, and help to improve urban resilience and scientific decision-making capability.

Keywords: SA-Informer Model; Risk Assessment; Precipitation Prediction; Ternary Coupling Coordination Degree Model.

1. Introduction

Increasing global climate change and rapid urbanisation have led to a steady rise in urban flood risk. However, there are significant deficiencies in risk assessment and precipitation prediction: risk assessment relies on single domain data, which makes it difficult to dynamically quantify socio-economic vulnerability and disaster prevention and mitigation responses; in precipitation prediction, traditional time series models (e.g., ARIMA, LSTM) are unable to deal with the long-period dependence and fluctuation of extremes effectively, and are prone to overfitting problems [1]. In order to break through the above bottlenecks, this paper constructs a technical framework for urban flood risk assessment and precipitation prediction with SA-Informer model as the core.

First, this study integrates multi-source heterogeneous data indicators based on the P-S-R model [2], and reveals the significant impact of extreme precipitation events on urban flood risk through the AHP-CRITIC combined assignment algorithm and the improved ternary coupled coordination degree model [3]. Using a stepwise regression model, the indicators of the pressure subsystem P and the response subsystem R are analyzed, confirming that the amount of heavy precipitation is the key driving factor. The model provides data support and theoretical cornerstone for the construction of precipitation prediction models. In addition, this study innovatively constructs the SA-Informer model, globally optimizes the learning rate, input sequence length and other hyperparameters by simulated annealing algorithm [4], and integrates the long time series feature extraction capability of the Informer model to construct a multi-dimensional feature input space including historical precipitation, meteorological elements [5-6], etc., which effectively captures the nonlinear patterns and seasonal fluctuations of precipitation data. Through the deep synergy between risk assessment and prediction models, the framework forms a complete technical chain from disaster mechanism analysis, data feature mining to high-precision early warning, which can be generalized to other climate-sensitive cities in the fields of disaster prediction, risk management and sustainable development. First, based on the P-S-R model [2], this study integrates heterogeneous data indicators from multiple sources, and reveals the significant impact of extreme precipitation events on urban flood risk through the combined AHP-CRITIC assignment algorithm and the improved ternary coupled coordination degree model [3], confirming that the amount of heavy precipitation is the key driving factor. The model provides data support and theoretical cornerstone for the construction of precipitation prediction models. In addition, this study innovatively constructs the SA-Informer model, globally optimizes the learning rate, input sequence length and other hyperparameters by simulated annealing algorithm [4], and integrates the long time series feature extraction capability of the Informer model to construct a multi-dimensional feature input space including historical precipitation, meteorological elements [5-6], etc., which effectively captures the nonlinear patterns and seasonal fluctuations of precipitation data. Through the deep synergy between risk assessment and prediction models, the framework forms a complete technical chain from disaster mechanism analysis, data feature mining to high-precision early warning, which can be generalized to other climate-sensitive cities in the fields of disaster prediction, risk management and sustainable development.

2. Urban Flood Risk Assessment and Analysis

2.1. Construction of Indicator System

The PSR model was proposed by Anthony Marcus Freund in 1979, and has been modified by relevant organisations to become a commonly used model for disaster risk assessment. In urban flooding, 'Pressure' refers to the external drivers that trigger the disaster, 'State' characterises the system characteristics at the time of the disaster, and 'Response' embodies the human countermeasures. As shown in Table.1, this study constructs a comprehensive assessment index system for urban flood disasters, in which the three major subsystems of pressure, state and response contain 4, 6 and 4 indicators, respectively.

Table 1. Indicator system for urban flood risk assessment

Target Layer	Normative Layer	Indicator Layer	Description of Indicators
Flood risk in Hefei	P	P1(-)	Intense precipitation R95p
		P2(+)	GDP per capita
		P3(+)	Population density
		P4(+)	Urban built-up area
	S	S1(-)	Runoff depth
		S2(-)	Population affected
		S3(-)	Area affected by crops
		S4(-)	Damage to sluice gates
		S5(-)	Length of damaged embankment
		S6(-)	Direct economic losses
	R	R1(+)	Length of drainage pipes
		R2(+)	Large reservoir storage
		R3(+)	Soil erosion control area, flood control area
		R4(+)	Area of green space
Note: (+) represents positive indicators and (-) represents negative indicators.			

2.2. Calculation of Indicator Weights

In this study, for the indicator system of the three subsystems, the hierarchical analysis method was used to systematically analyse their subjective weights, while the CRITIC method was applied to quantitatively assess their objective weights. On this basis, the weights obtained from the two methods are integrated through the linear assignment method to finally form the comprehensive weights. Among them, the difference coefficient method is used to determine the distribution ratio of the combined weights to ensure the scientificity and rationality of the weight distribution.

Let the combination weight vector $w = \alpha w' + \beta w''$, α, β be the to-be-determined coefficients for assigning weights to the main objective combinations, satisfy: $\alpha, \beta \geq 0, \alpha + \beta = 1$.

$$\begin{cases} \alpha = \dfrac{n}{n-1} \cdot T' \\ T' = \dfrac{2}{n} \sum_{i=1}^{n} (i \cdot p_i) - \dfrac{n+1}{n} \end{cases} \tag{1}$$

Where T' is the coefficient of variation for each component of w' and $p_1, p_2, \cdots, p_n$ is the rearrangement of the components of the subjective weight vector from smallest to largest. After obtaining α, β can be calculated by combining $\alpha + \beta = 1$, and then substituting into $w = \alpha w' + \beta w''$ to obtain the combined weight vector $w = (w_1, w_2, \cdots, w_n)^T$ for each attribute.

Finally, a composite score for each subsystem was calculated:

$$\begin{cases} P_i = \sum_{j=1}^{n} w_i \cdot P_{ij} \\ S_i = \sum_{j=1}^{n} w_i \cdot S_{ij} \\ R_i = \sum_{j=1}^{n} w_i \cdot R_{ij} \end{cases} \tag{2}$$

Where P_i, S_i, R_i represents the composite score of each subsystem in year i, and represents the normalised value of each evaluation indicator, respectively. The comprehensive weighting system fully integrates the advantages of subjective judgement and objective data, providing a scientific and comprehensive quantitative basis for urban flood risk assessment.

2.3. Analysis of Coupled Coordination Relationships between Subsystems

1) Improved coupled coordination degree model and flood risk rating scale

The structure of the coupling coordination degree model consists of three parts: coupling degree, coordination degree and coupling coordination degree. Coupling degree C reflects the strength of inter-system interactions, coordination degree T measures the consistency of the developmental dynamics, and coupling coordination degree D comprehensively evaluates the level of system synergy, providing a scientific basis for the optimisation of complex systems.

In order to solve the triple limitations of the traditional model of 'coupling' and 'coordination' in terms of ill-defined semantic boundaries, fundamental flaws in the boundary value processing mechanism, and structural flaws in the differentiation effectiveness, this paper adopts a new coupling coordination improvement model with three core assumptions: an ideal coupling state, a maximum deviation boundary, and the independence of coupling and coordination. The k-element improved coupling coordination degree formula is as follows:

$$\begin{cases} C = 1 - 2\sqrt{\dfrac{k \cdot \sum_{i=1}^{k} U_i^2 - \left(\sum_{i=1}^{k} U_i\right)^2}{k^2}} \\ T = \sum_{i=1}^{k} a_i \times U_i, \quad \sum_{i=1}^{k} a_i = 1 \\ D = \sqrt{C \times T} \end{cases} \tag{3}$$

Where k is the number of subsystems and U_i is the subsystem score.

In this study, based on the coupled coordination degree model and with reference to the relevant critical thresholds, the risk level of urban flooding is optimally adjusted and a five-level risk level system is constructed. This system not only retains the original data distribution characteristics, but also reduces the management complexity and effectively supports the formulation of hierarchical control strategies.

Table 2. Five-level risk rating scale

Risk Level	Range	Risk Description
Extremely High Risk	[0,0.359)	Severe system dysfunction and extremely high risk
High Risk	[0.359,0.430)	Inadequate multi-system coordination and high risk
Medium Risk	[0.430,0.495)	Some systems are weak and need improvements
Low Risk	[0.495,0.578)	Initial coordination, but potential pitfalls remain
Extremely Low Risk	[0.578,1]	Highly synergistic system with manageable flood risk

By checking the calculated coupling coordination degree of the three P-S-R subsystems against the risk grading table of this study, the risk level of urban flooding can be visually assessed, and the results of urban flooding risk assessment can be obtained.

2) Analysis of the degree of coupling coordination

In calculating the degree of coordination, the state subsystem (S) reflects the impacts of flooding, the pressure subsystem (P) incorporates drivers such as extreme precipitation and urbanisation, and the response subsystem (R) captures the capacity and adaptability of the city's governance. P and R form a dynamic equilibrium: P increases risk, and R mitigates risk. Treating P and R as equally important reflects symmetry and ensures that the model fully reflects the dynamic equilibrium. To highlight the centrality of S and balance the contributions of P and R, this study assigns the weights of P, S, and R as $\frac{1}{5},\frac{3}{5},\frac{1}{5}$, in order to more scientifically assess the overall coordination and risk of urban flooding systems.

The ternary coupling coordination degree D values between the pressure subsystem P, state subsystem S and response subsystem R for the years 2006-2022 were calculated by the improved coupling coordination degree model, as shown in Figure 1. The D-value shows the characteristics of 'stepwise optimisation-sudden fluctuation' during the period of 2006-2022: in the first period (2006-2012), the D-value continues to climb from 0.544 to 0.724, entering into a highly coordinated steady state (2013-2019, excluding 2016); it plummets to 0.42 in 2020 but rapidly recovers to 0.85+ in the following year, reflecting the system's resilience. It is noteworthy that the D-value anomaly dropped to 0.417 in 2016, forming a faulty contrast with the high coordination period, indicating that the P-S-R system experienced an equal degree of dissonanc in these two years.

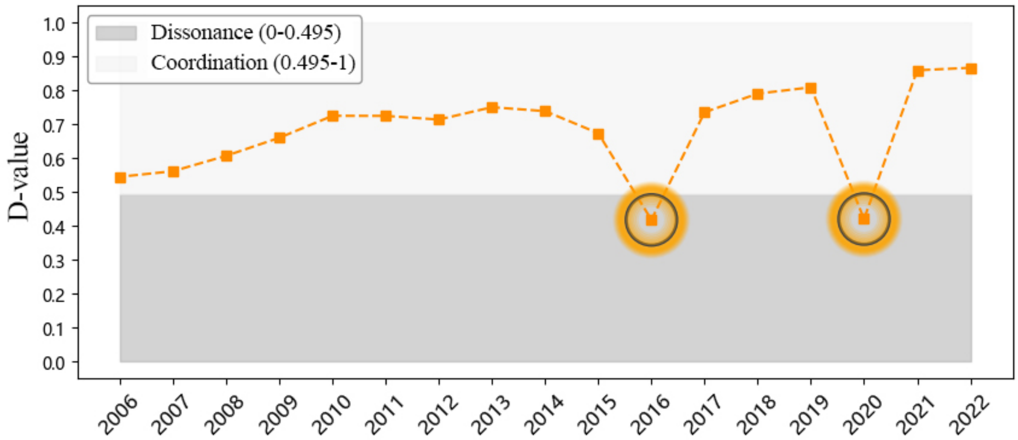

Figure 1. P-S-R system coupling coordination degree

In order to further identify the key factors leading to high risk of urban flooding, this study analysed the influence of the indicators of the pressure subsystem P and the response subsystem R on the degree of coupling coordination with the help of a stepwise regression model ($F = 28.373, P < 0.001$). The results show that the strong precipitation R95p and the length of the drainage pipe have a significant effect on the state coupling coordination degree, with the contribution of the strong precipitation R95p being more prominent. This leads to the conclusion that heavy precipitation is the main driver that triggers high risk of urban flooding. This conclusion highlights the central role of extreme precipitation events in flood risk formation, and suggests that the optimisation of urban drainage systems still needs to be further enhanced.

3. Precipitation Prediction based on SA-Informer

It has been sufficiently discussed that heavy precipitation is one of the key factors affecting the degree of coordination of the system coupling, and it has a significant impact on urban flood risk. Therefore, choosing an appropriate precipitation prediction model to predict future precipitation is of great practical importance and urgency for formulating effective defence strategies in advance and mitigating future urban flood risks.

3.1. SA-Informer Model

In this paper, the simulated annealing algorithm is innovatively applied to the optimisation process of the model. The specific optimisation ideas are as follows:

Step 1: Hyperparameter space definition and neighbourhood perturbation

1) Learning rate: logarithmic spatial perturbation is used to avoid the search blindness of traditional linear perturbation in the region of very small values by multiplying by a random factor (0.1~10.0) and limiting the range (0.00001~0.01).

2) Penalty factor: Balance false alarms and missed alarms by uniform perturbation (±1.0) over the range.

3) Input sequence length: based on the periodicity of the time series, it is randomly adjusted in integer steps in the interval [10,90] to capture precipitation patterns at different time scales.

4) Number of hidden layer units: adjusted in multiples of 16 (32~256) to ensure that matrix operations are aligned to GPU memory bandwidth and improve computational efficiency.

Step 2: Annealing process and loss function design

The initial high temperature allows a high probability of accepting inferior solutions to avoid falling into a local optimum; the temperature is gradually reduced by exponential cooling so that the model gradually converges to the global optimum solution. The loss function integrates the three valuation metrics RMSE, SMAPE, and RMSLE to penalise the underestimation of extreme precipitation events. Among them, SMAPE strengthens the sensitivity to small precipitation errors, and RMSLE suppresses the influence of outliers.

Step 3: Optimising processes

Set the baseline parameters as a starting point for the search. Generate new parameter combinations at each iteration and compute the validation set composite loss. Accept the inferior solution with probability $P = \exp\left(-\frac{\Delta Loss}{T}\right)$. If accepted, update the current solution; if not, keep the original solution unchanged. Record the historical optimal parameters and finally return the optimal parameter combination for full training.

In this study, the simulated annealing algorithm was experimentally implemented using Python programming to determine the optimal combination of learning rate, penalty coefficient, input sequence length, and number of hidden layer units as [0.000118, 2.7, 71, 96]. And the current optimal combination of parameters is imported into the Informer model, and the test set results are obtained as shown in Figure 2.

Figure 2 demonstrates the performance of the Informer model after incorporating the simulated annealing algorithm in the day-by-day precipitation prediction in Hefei. In general, the fluctuation trends of the predicted and actual values are more consistent, especially in the ups and downs of precipitation, which shows a better match and captures the dynamic changes of precipitation better. At most time steps, the predicted and actual values are close to each other, but there are still some deviations at the sudden precipitation change points.

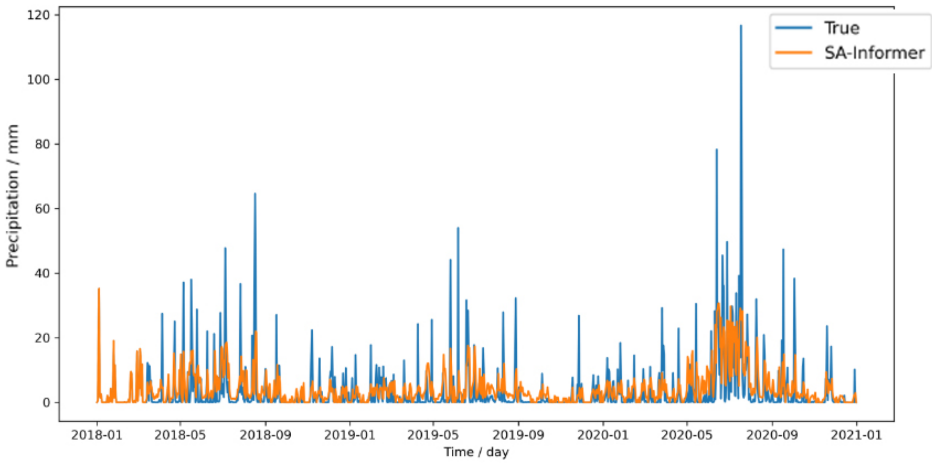

Figure 2. Prediction results of SA-Informer model on the test set

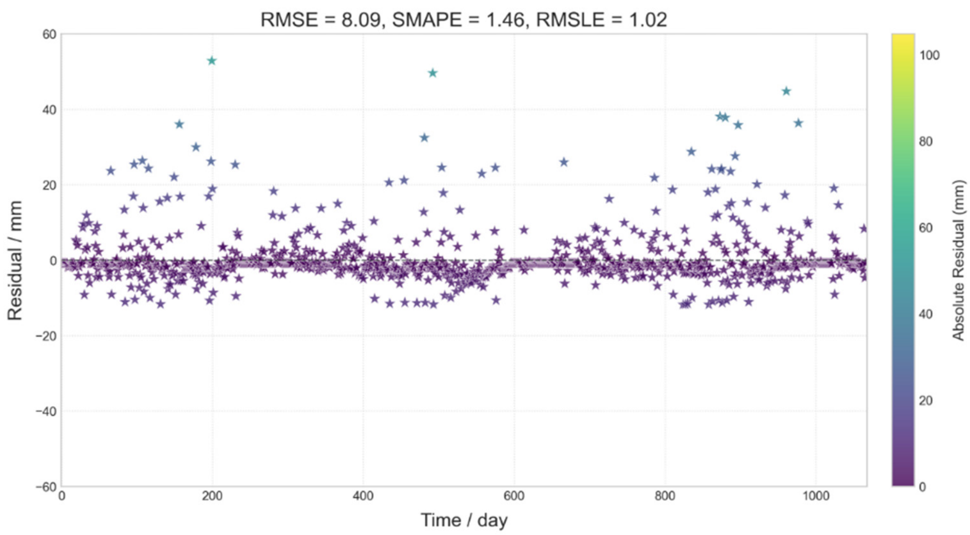

Figure 3. Test set residuals and evaluation results

Figure 3 demonstrates the results of the residual analysis of the Informer model after incorporating the simulated annealing algorithm. It can be seen that the residuals are mostly concentrated near 0, indicating that the predicted values of the model are closer to the actual values and the prediction accuracy is higher. At the same time, the distribution of the residuals on the time axis is relatively stable, with no obvious trend or cyclical changes, indicating that the model's predictive performance is more consistent in different time periods. The values of the three indicators RMSE, SMAPE and RMSLE are given in the figure as 7.74, 1.24 and 1.00, respectively, which are low, further indicating the accuracy and stability of the model prediction.

Figure 4 shows the results of the evaluation metrics of the Informer model incorporating the simulated annealing algorithm on the validation set, presenting the relationship between the three metrics of RMSE, SMAPE and RMSLE through a 3D scatter plot with the intensity of RMSLE indicated by colour. Observing the distribution of data points in the graph, it can be found that they are relatively concentrated in the region of lower RMSE and SMAPE values, while the RMSLE values are also relatively low and stable, indicating that the model performs better in terms of prediction accuracy and stability. This indicates that the optimised model of the simulated annealing algorithm has smaller errors on different validation set samples, and the prediction results are closer to the actual values, while the sensitivity to outliers is reduced and the prediction performance is more reliable.

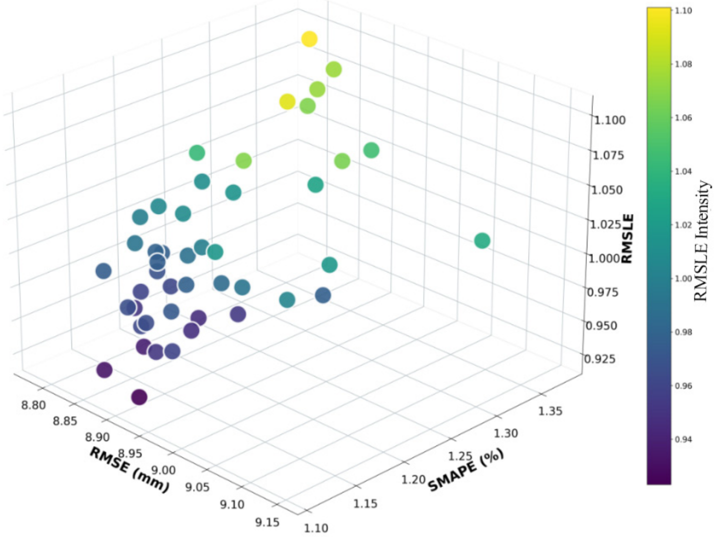

Figure 4. Three-dimensional scatterplot of validation set assessment metrics

Table.3 presents the performance evaluation of the results of fitting the day-by-day precipitation of Hefei in the comparison test. It can be seen that the Informer model fits the day-by-day precipitation in Hefei with an RMSE of 8.09, a SMAPE of 1.46, and an RMSLE of 1.02, while the corresponding RMSE of the SA-Informer model decreases to 7.74, the SMAPE decreases to 1.24, and the RMSLE decreases to 1.00. It can be seen that the SA-Informer model reduces the deviation between the true value and the fitted value by 4.33%, i.e., the model accuracy is improved by 4.33%, and its trained model is more accurate and the predicted value is closer to the true value.

Table 3. Comparative experimental performance evaluation results

	RMSE	SMAPE	RMSLE
Informer	8.09	1.46	1.02
SA—Informer	7.74	1.24	1.00

3.2. Analysis of Model Results

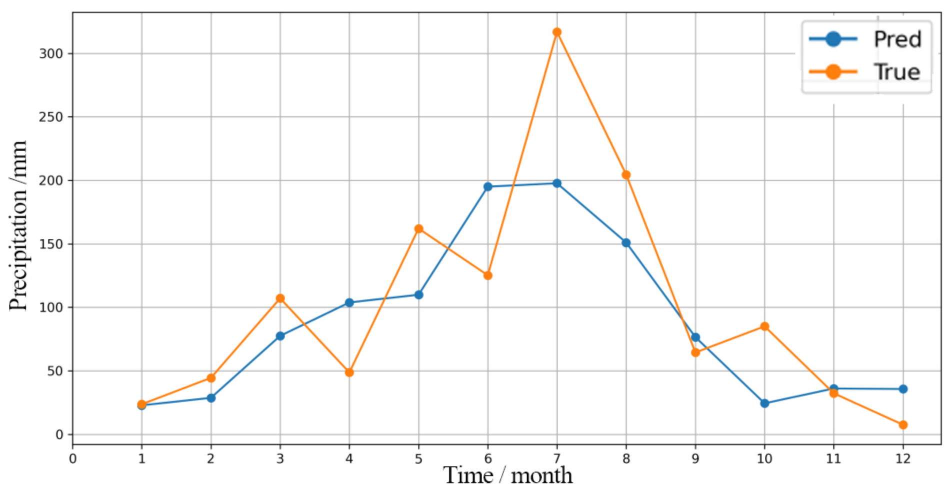

Figure 5. SA-Informer model predictions of month-by-month precipitation for Hefei in 2021

The Informer model incorporating simulated annealing algorithm is used to predict the month-by-month precipitation in Hefei City in 2021, and the prediction results are shown in Figure 5.

From the figure, it can be seen that the SA-Informer model predicts the month-by-month precipitation in Hefei City in 2021 with satisfactory results, and the overall trend of the predicted and real values is basically the same, which successfully captures the seasonal change pattern of precipitation. It is further shown that the SA-Informer model is able to reflect the precipitation characteristics of Hefei City more accurately after optimising the combination of parameters, which provides a reliable prediction basis for subsequent meteorological warnings and flood risk assessment.

In view of this, other researchers can choose the SA-Informer model to predict the future precipitation data of the city, and then further calculate the R95p value for each year through the precipitation data obtained from the prediction, and calculate the D value of the degree of coupling coordination based on the R95p in combination with the development trend of the future pressure, state, and response indicators. Eventually, the risk level of urban flooding can be accurately assessed against the risk level classification criteria (Table.2).

4. Conclusion

The technical framework of urban flood risk assessment and precipitation forecasting constructed in this paper with the SA-Informer model as the core effectively breaks through the bottleneck of the existing research in the integration of multi-source data and long time series forecasting. First, this study constructs a risk assessment system based on the P-S-R model integrating heterogeneous data from multiple sources, and analyzes the impacts of the pressure subsystem and response subsystem indicators on the coupling degree of coordination through the combination of the AHP-CRITIC assignment and the improvement of the ternary coupling degree of coordination model, combined with the stepwise regression model, which confirms that the heavy precipitation as a core driver is significantly related to the increase in risk level, and lays the foundation for the construction of the prediction model. The construction of the prediction model lays the foundation of causal logic. On this basis, the study innovatively proposes the SA-Informer model, which realizes the adaptive optimization of key parameters such as the learning rate and input sequence length by simulating the global search mechanism of the annealing algorithm. Combined with the seven-dimensional feature input space, the prediction accuracy of the model is improved by 4.33%, which significantly enhances the ability to capture the nonlinear pattern and seasonal fluctuation of precipitation data.

The framework forms a complete technology chain through the in-depth synergy of risk assessment and prediction models, which has been empirically verified to effectively improve the comprehensiveness of flood risk assessment and the accuracy of precipitation prediction, and provide a scientific and reliable solution for the prevention and control of urban flooding disasters. Future research could explore the dynamic fusion mechanism with real-time monitoring data to improve the timeliness of extreme weather warning and the accuracy of prevention and control.

References

[1] Yang Haimin, Pan Zhisong, Bai Wei. A review of time series forecasting methods [J]. Computer Science, 2019, 46 (01): 21-28.

[2] Yu Guangyong, Gu Yangyang, Zhang Minxia, et al. Dynamic analysis of ecological carrying capacity in Quanzhou City, Fujian Province based on the "pressure-state-response" model [J]. Environmental Ecology, 2024, 6 (05): 1-6.

[3] Fan Dongjun, Ke Ran, Cao Runmin. Modification and improvement of the coupled coordination degree model [J]. Statistics and Decision Making, 2024, 40 (22): 41-46.

[4] Li Yanbo, Li Ruochen, Shi Bo, et al. Energy-efficient optimisation of self-consistent energy system in highway service area based on improved simulated annealing genetic algorithm [J]. Journal of Xi'an Jiaotong University, 2024, 58 (01): 197-207+216.

[5] Liu Hongxiao, Xiang Mian, Zhou Bintao, et al. Long-series time-series power load forecasting based on Informer [J]. Journal of Hubei University for Nationalities (Natural Science Edition), 2021, 39 (03): 326-331.

[6] Ma Guizhao, Zhang Yuhang, Wang Lingfei. TimeGAN-Informer Long-Term Airport Visibility Prediction [J]. Journal of Safety and Environment, 2024, 24 (07): 2517-2527.