Seq2Seq-DDPM: A novel interval flood forecasting model based on Sequence-to-Sequence Denoising Diffusion Probabilistic Model

Chang Chen[1] and Dawei Zhang[2]
State Key Laboratory of Water Cycle and Water Security, China Institute of Water Resources and Hydropower Research, Beijing 100038, China[1]
Research Center on Flood & Drought Disaster Prevention and Reduction, Ministry of Water Resources, Beijing 100038, China[1]
E-mail: 2234104368@qq.com

State Key Laboratory of Water Cycle and Water Security, China Institute of Water Resources and Hydropower Research, Beijing 100038, China[2]
Research Center on Flood & Drought Disaster Prevention and Reduction, Ministry of Water Resources, Beijing 100038, China[2]
E-mail: zhangdw@iwhr.com

ABSTRACT

Accurate flood forecasting performs a vital function in safeguarding basin water security. In recent years, neural network-based hydrological models have shown strong predictive power. However, uncertainties in data, parameters, and structure lead to variability in forecasts. Thus, an innovative flood forecasting model, Seq2Seq-DDPM, is proposed by the integration of the Sequence-to-Sequence (Seq2Seq) framework and the Denoising Diffusion Probabilistic Model (DDPM). It is specifically designed to achieve probabilistic forecasting of streamflow processes, quantifying hydrological uncertainties. It employs the Bidirectional Gate Recurrent Unit (BiGRU) as the Encoder and the Gate Recurrent Unit (GRU) as the Decoder, with the diffusion model serving as the bridge connecting the two. This architecture encapsulates the uncertainty of hydrological processes within the hidden state. Comparative experiments are carried out in the Xiapu River Basin to verify the effectiveness of probabilistic forecasting models. The proposed Seq2Seq-DDPM is compared against various mainstream models, including the Probabilistic Forecasting with Autoregressive Recurrent Networks (DeepAR), the Lower and Upper Bound Estimation (LUBE), and the Quantile Regression (QR) method. The findings confirm that the proposed Seq2Seq-DDPM model achieves the highest Prediction Interval Coverage Probability (PICP) value and the lowest Prediction Interval Normalized Averaged Width (PINAW) value among the compared methods. The forecast intervals reliably encompass most of the observed streamflow values, indicating its superior probabilistic forecasting capability. These outcomes suggest that the model can provide effective risk information for flood control scheduling decisions.

KEYWORDS: Flood probabilistic forecasting; Seq2seq; Diffusion model; Neural networks;

1 INTRODUCTION

Flood forecasting constitutes a crucial non-engineering measure, offering a scientific foundation in basin flood prevention and disaster mitigation, while also offering essential support for the optimal management of water resources. With the rapid advancement of artificial intelligence technology, deep learning has demonstrated great potential in flood forecasting. Among these, recurrent neural networks have captured notable focus in flood forecasting, due to their proven efficacy in sequence prediction. Representative architectures, typified by the Long Short Term Memory (LSTM) and its variants, have been extensively applied in hydrological modelling. Liu et al., (2023) constructed the LSTM model and applied it to forecast streamflow in the area between the Xiangtan Station and Changsha Station, thereby demonstrating its capability in flood early warning. Yan et al., (2023) proposed a novel flood forecasting model that integrates the simulated annealing (SA) algorithm and the Gate Recurrent Unit (GRU), enabling daily-scale flood forecasting in the central urban area of Hefei. Boughale et al., (2024) utilized the Historical Daily Weather dataset from the Australian Commonwealth Office of Meteorology to test the Bidirectional Long Short Term Memory (BiLSTM). The findings reveal that the BiLSTM model consistently achieved the best forecasts. Fan and Chang, (2024) used the Bidirectional Gate Recurrent Unit (BiGRU) to predict the streamflow and water level in the Xiaoqing River Basin, maintaining a high level of prediction accuracy.

It is important to note that the aforementioned studies are all deterministic forecasts. However, due to the complexity of hydrological processes, flood forecasting inevitably involves uncertainties related to input data, model structures, and model parameters (Matthews et al., 2021). These uncertainties affect the reliability of prediction results, leading to suboptimal robustness in deterministic forecasts. Currently, the deep integration of advanced uncertainty modelling techniques with deep learning frameworks has developed into a vital direction in probabilistic flood forecasting. Chang et al., (2024) created a method based on the Lower and Upper Bound Estimation (LUBE), which illustrated strong performance in reducing forecast uncertainty in the Yalong River Basin. Weng et al., (2023) trained the Quantile Regression (QR) method integrated with LSTM networks (QRLSTM) using both real and synthetic flood sequences in the flood interval prediction field. Xie et al., (2025) designed a modelling framework based on the Probabilistic Forecasting with Autoregressive Recurrent Networks (DeepAR) to better capture the variability of flow uncertainty in two basins.

This study selects the Xiapu River Basin above the Dawuchang hydrological station in Xianning City as the study area and constructs the Seq2Seq-DDPM probabilistic flood forecasting model by integrating the Sequence-to-Sequence (Seq2Seq) framework with the Denoising Diffusion Probabilistic Model (DDPM). The experimental results of Seq2Seq-DDPM are compared and analysed against those of commonly used methods, thereby providing a technical reference for future probabilistic flood forecasting research.

2 METHOD

An innovative flood forecasting model, Seq2Seq-DDPM, is proposed by the integration of the Sequence-to-Sequence (Seq2Seq) framework and the Denoising Diffusion Probabilistic Model (DDPM), as shown in Figure 1. The model combines the Seq2Seq architecture with the DDPM to achieve probabilistic forecasting of streamflow processes. This innovative framework utilizes diffusion processes to encapsulate uncertainty within hidden states, not only providing uncertainty estimates in hydrological models but also significantly reducing computational cost.

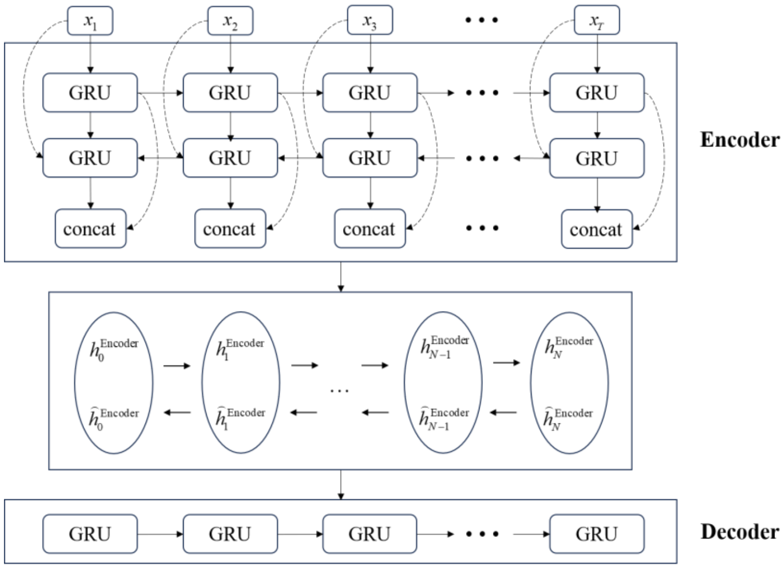

Figure 1: The framework of Seq2Seq-DDPM

The sequence-to-sequence (Seq2Seq, also known as Encoder-Decoder) framework proposed by Sutskever et al., (2014) leverages the information transmission mechanism of the Encoder-Decoder, offering significant advantages in time-series processing and forecasting. Notably, its application in hydrology can effectively capture the dependencies in rainfall-runoff relationships, enhancing the predictive accuracy (Kao et al., 2020; Xiang et al., 2020).

In the Seq2Seq-DDPM model, the Encoder processes input sequences in both forward and backward directions using the BiGRU architecture, enabling the extraction of bidirectional features from flood time series. This bidirectional mechanism enables the retention and more comprehensive exploitation of information during flood peaks. The gating mechanism of the Gate Recurrent Unit (GRU) in the Decoder facilitates the selective retention and updating of information, thereby enhancing the accuracy of the decoding process. The structure of GRU is illustrated in Figure 2, which includes components such as update gates z_t and reset gates r_t to control historical information x_t flow. BiGRU extends the GRU architecture by concatenating the hidden states from both the forward direction $\vec{h}_t$ and the backward direction $\overleftarrow{h}_t$, as defined in Eqs. (1) to (3).

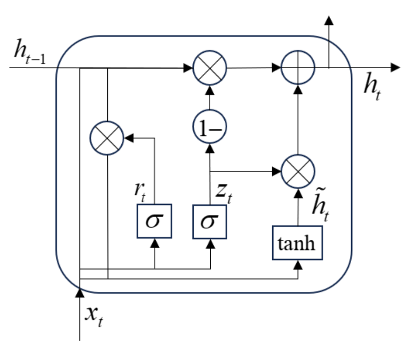

Figure 2: The framework of GRU

$$\vec{h}_t = \begin{cases} \vec{z}_t = \sigma\left(\vec{W}_z \vec{x}_t + \vec{U}_z \vec{h}_{t-1} + \vec{b}_z\right) \\ \vec{r}_t = \sigma\left(\vec{W}_r \vec{x}_t + \vec{U}_r \vec{h}_{t-1} + \vec{b}_r\right) \\ \tilde{h}_t = \tanh\left(\vec{W} \vec{x}_t + \vec{U}\left(\vec{r}_t \odot \vec{h}_{t-1}\right)\right) \\ \vec{h}_t = \left(1 - \vec{z}_t\right) \odot \vec{h}_{t-1} + \vec{z}_t \odot \tilde{h}_t \end{cases} \tag{1}$$

$$\overleftarrow{h}_t = \begin{cases} \overleftarrow{z}_t = \sigma\left(\overleftarrow{W}_z \overleftarrow{x}_t + \overleftarrow{U}_z \overleftarrow{h}_{t-1} + \overleftarrow{b}_z\right) \\ \overleftarrow{r}_t = \sigma\left(\overleftarrow{W}_r \overleftarrow{x}_t + \overleftarrow{U}_r \overleftarrow{h}_{t-1} + \overleftarrow{b}_r\right) \\ \tilde{h}_t = \tanh\left(\overleftarrow{W} \overleftarrow{x}_t + \overleftarrow{U}\left(\overleftarrow{r}_t \odot \overleftarrow{h}_{t-1}\right)\right) \\ \overleftarrow{h}_t = \left(1 - \overleftarrow{z}_t\right) \odot \overleftarrow{h}_{t-1} + \overleftarrow{z}_t \odot \tilde{h}_t \end{cases} \tag{2}$$

$$h_t = \left[\vec{h}_t \oplus \overleftarrow{h}_t\right] \tag{3}$$

where $\rightarrow$ and $\leftarrow$ are the forward and backward processes, respectively. $\tilde{h}_t$ is the hidden input activation vector. b_z and b_r indicate the biases, respectively. W_z, U_z, W_r, U_r and U refer to the weights, respectively. $\sigma(\cdot)$ and $\tanh(\cdot)$ are activation functions, respectively. $\odot$ is an element-wise multiplication. $\oplus$ illustrates an element-wise sum.

Serving as a crucial bridge between the Encoder and the Decoder, the diffusion model provides a probabilistic representation of the latent state space. In hydrological probabilistic forecasting, its core value is to offer a probabilistic modelling framework that generates runoff's conditional probability distributions through the transformation of latent variables using a Markov chain. This approach breaks away from the single-output mode of traditional deterministic forecasting.

Diffusion models consist of forward and reverse processes. The former develops a Markov chain by adding noise $\varepsilon \sim \mathrm{N}\left(0, \mathbf{I}\right)$. The latent variable h_N is shown as Eq. (4), after N steps.

$$h_N = \sqrt{\bar{\alpha}_N} h_0^{\text{Encoder}} + \sqrt{1 - \bar{\alpha}_N} \varepsilon \tag{4}$$

where h_0^{Encoder} is the hidden state output from the Encoder. α_N weights the latent variable and the noise and $\bar{\alpha}_N = \prod_{n=1}^{N} \alpha_n$. The desired distribution of the hidden state is shown in Eq. (5).

$$\begin{cases} q\left(h_1, h_2, \cdots, h_N \middle| h_0\right) = \prod_{n=1}^{N} q\left(h_n \middle| h_{n-1}\right) \\ q\left(h_n \middle| h_{n-1}\right) = \mathrm{N}\left(h_n; \sqrt{1 - \beta_n} h_{n-1}, \beta_n \mathbf{I}\right) \end{cases} \tag{5}$$

where $\beta_n = 1 - \alpha_n$. The reverse diffusion process is formulated as Eq. (6).

$$\begin{cases} p_\theta\left(h_0,h_1,\cdots,h_{n-1}\middle|h_N\right)=p\left(h_N\right)\prod_{n=1}^{N}p_\theta\left(h_{n-1}\middle|h_n\right) \\ p_\theta\left(h_{n-1}\middle|h_n\right)=\mathrm{N}\left(h_{n-1};\mu_\theta\left(h_n,n\right),\tilde{\beta}_n\mathbf{I}\right) \\ \tilde{\beta}_n=\left(1-\bar{\alpha}_{n-1}\right)\beta_n/1-\bar{\alpha}_n \end{cases} \tag{6}$$

where $p\left(h_N\right)\sim\mathrm{N}\left(0,\mathbf{I}\right)$ and $\tilde{\beta}_n=\left(1-\bar{\alpha}_{n-1}\right)\beta_n/1-\bar{\alpha}_n$. $\mu_\theta\left(h_n,n\right)$ denotes the mean value.

3 STUDY CASE

The study area selects for this paper is the Xiapu River Basin above the Dawuchang hydrological station in Xianning City, China, covering an area of 112.5 km^2, as shown in Figure 3. There are three rainfall stations within the catchment: Sanbao Station, Linshang Station, and Huangjin Station. The Dawuchang Hydrological Station is located at the catchment outlet (Zhang et al., 2015).

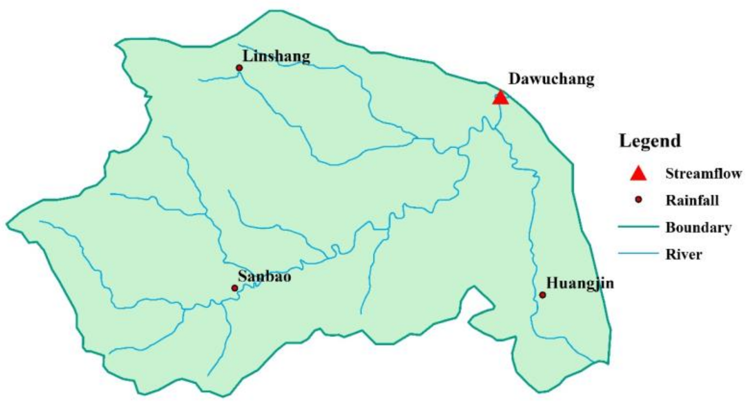

Figure 3: Study area

For probability prediction, two evaluation metrics are used: the Prediction Interval Coverage Probability (PICP), which assesses the percentage of actual values that lie in the forecast interval, and the Prediction Interval Normalized Averaged Width (PINAW), which quantifies the width of the interval relative to the range of observed data. These are defined in Eqs. (7) to (9). The higher the likelihood that the observed value lies in the forecast interval, the greater the PICP value and the better the prediction performance. Conversely, the larger the width of the forecast interval, the bigger the PINAW value and the poorer the predictive capability.

$$c_t=\begin{cases}1 & Q_t^{\mathrm{obs}}\in\left[L_t,U_t\right] \\ 0 & Q_t^{\mathrm{obs}}\notin\left[L_t,U_t\right]\end{cases} \tag{7}$$

$$\mathrm{PICP}=\frac{1}{M}\sum_{t=1}^{M}c_t \tag{8}$$

$$\mathrm{PINAW}=\frac{1}{MR}\sum_{t=1}^{M}\left(U_t-L_t\right) \tag{9}$$

where c_t denotes a Boolean variable. Q_t^{obs} is the observed flows. R indicates the range of the variable. M represents the count of samples. U_t and L_t denote the upper and lower limits.

4 RESULTS

This study reports the evaluation metrics for the 90% confidence interval with different lead times. Figure 4(a) and Figure 4(b) present, respectively, statistical heatmaps of the mean values for the PICP and PINAW evaluation indices in the Xiapu River Basin. These experiments examine the reliability and concentration of probabilistic forecasts from different perspectives.

When the lead time is 1h-3h, the PICP values for Seq2Seq-DDPM are 90.42%-93.68%. The PINAW values for Seq2Seq-DDPM are 0.085-0.098. The PICP values for QR, LUBE, and Seq2Seq-DDPM all exceed or approach the 90% confidence level. However, LUBE has an excessively high PINAW value, suggesting that its interval forecast results are relatively unreliable. Meanwhile, the Seq2Seq-DDPM model exhibits an average PICP value that is 0.7% higher than QR's, while its PINAW value decreases by an average of 0.003. In summary, the Seq2Seq-DDPM model delivers superior probabilistic forecasting performance with narrower intervals while keeping high coverage.

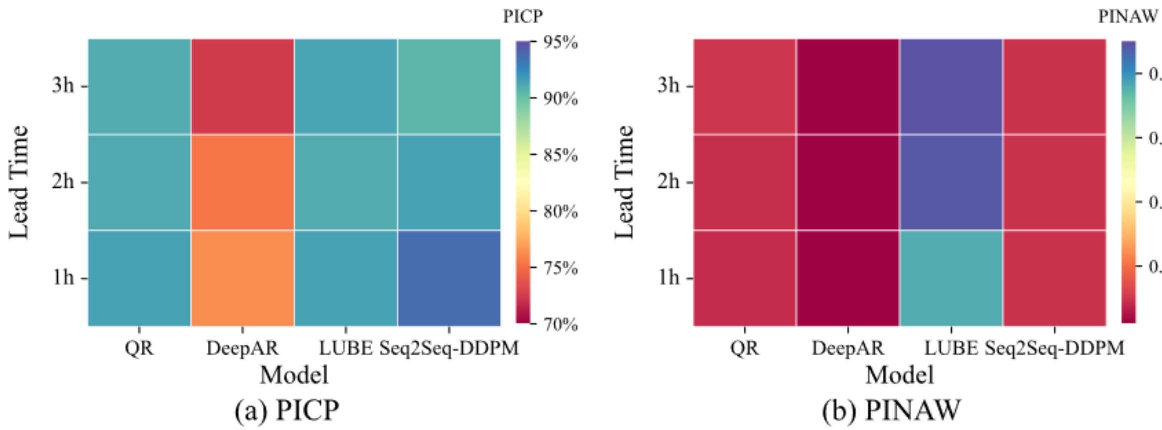

Figure 4: Statistical heatmaps of the average values for probabilistic evaluation metrics

Figure 5 and Figure 6 display the prediction intervals at the 90% confidence interval for Seq2Seq-DDPM, providing a visual analysis of the model's performance in quantifying prediction uncertainty. Using the No. 020513 and No. 040429 flood events at 1h forecast horizon as examples, the Seq2Seq-DDPM model covers most of the observed flow, indicating reasonable reliability of the prediction intervals. However, for the No. 020513 flood event, the Seq2Seq-DDPM model slightly underestimates the peak flows. Enhancing the model's performance in predicting flood peaks is a potential area for future work.

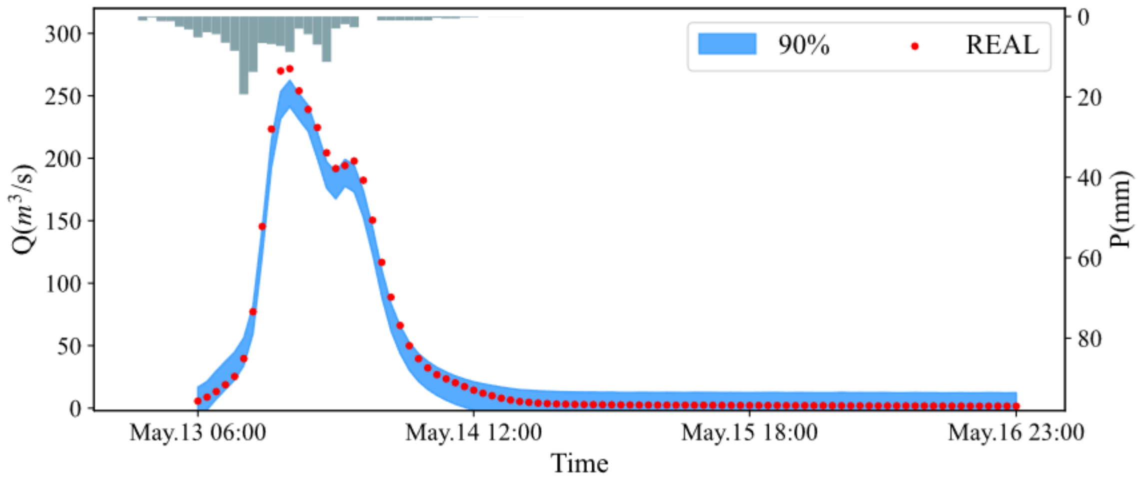

Figure 5: The forecast result of the No. 020513 flood event

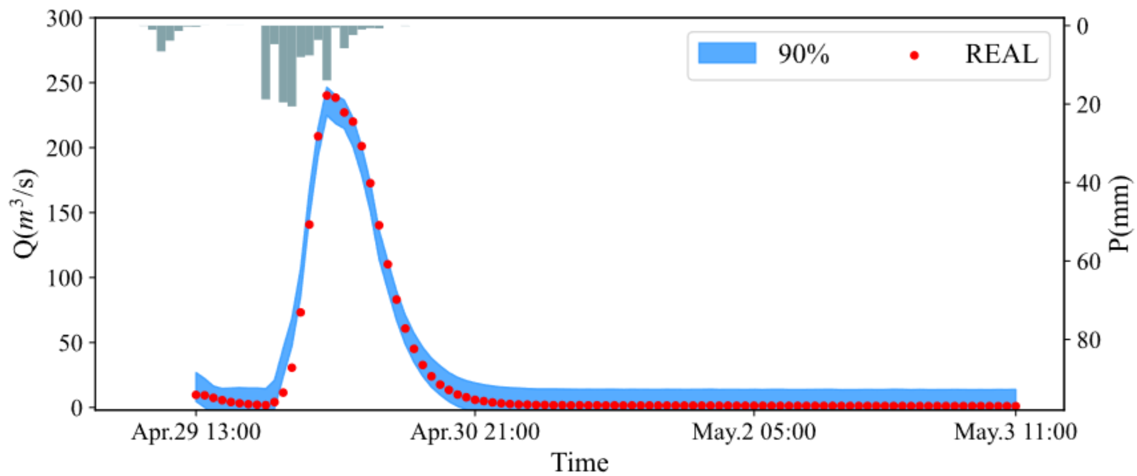

Figure 6: The forecast result of the No. 040429 flood event

5 CONCLUSION

This paper presents a flood probability model based on the diffusion model and the Seq2Seq architecture, termed Seq2Seq-DDPM. It employs the BiGRU as the Encoder and the GRU as the Decoder, with the diffusion model serving as the bridge connecting the two. This architecture encapsulates the uncertainty of hydrological processes within the hidden state, quantifying hydrological uncertainties. The Seq2Seq-DDPM uses observed rainfall and streamflow as inputs to probabilistically forecast future streamflow.

Experimental results demonstrate that the Seq2Seq-DDPM model obtains the highest PICP and the lowest PINAW, which shows that it can create a narrow interval while maintaining high coverage. For model-fitting performance, the Seq2Seq-DDPM model covers most of the observed flow, indicating reasonable reliability of the prediction intervals.

Meanwhile, Seq2Seq-DDPM still presents opportunities for further improvement. The study observes that, in certain scenarios, the model underestimates flood peaks. Consequently, future work may consider developing a novel loss function that effectively reduces peak errors to further enhance the model's predictive performance.

6 ACKNOWLEDGEMENTS

The study was supported by the National Key Research and Development Program of China (Grant No. 2024YFC3214802) and the IWHR Research & Development Support Program (WH0145B022021).

REFERENCES

Boughale R., Zrelli A. and Ezzedine T. (2024). Enhancing Flood Forecasting with BiLSTM Networks, in "2024 IEEE/ACS 21st International Conference on Computer Systems and Applications (AICCSA),"IEEE, 2024, pp. 1–6.

Chang X., Guo J., Qin H., Huang J., Wang X. and Ren P. (2024). Single-Objective and Multi-Objective Flood Interval Forecasting Considering Interval Fitting Coefficients. Water Resour. Manag., 38, 3953–3972.

Fan W. and Chang C. (2024). Study Of Flood Forecasting Based on Recurrent Neural Network for Urban River in The Piedmont Plain, in "15th International Conference on Hydroinformatics,"2024, p. 289.

Kao I.-F., Zhou Y., Chang L.-C. and Chang F.-J. (2020). Exploring a Long Short-Term Memory based Encoder-Decoder framework for multi-step-ahead flood forecasting. J. Hydrol., 583, 124631.

Liu Y., Yang Y., Chin R.J., Wang Chucai and Wang Changshun (2023). Long Short-Term Memory (LSTM) Based Model for Flood Forecasting in Xiangjiang River. KSCE J. Civ. Eng., 27, 5030–5040.

Matthews G., Barnard C., Cloke H., Dance S.L., Jurlina T., Mazzetti C. and Prudhomme C. (2021). Evaluating the impact of post-processing medium-range ensemble streamflow forecasts from the European Flood Awareness System. Hydrol. Earth Syst. Sci. Discuss., 2021, 1–51.

Sutskever I., Vinyals O. and Le Q.V. (2014). Sequence to sequence learning with neural networks. Adv. Neural Inf. Process. Syst., 27.

Weng P., Tian Y., Liu Y. and Zheng Y. (2023). Time-series generative adversarial networks for flood forecasting. J. Hydrol., 622, 129702.

Xiang Z., Yan J. and Demir I. (2020). A Rainfall-Runoff Model With LSTM-Based Sequence-to-Sequence Learning. Water Resour. Res., 56, e2019WR025326.

Xie S., Wang D., Wang J., Yang C., Shen K., Jia B. and Cao H. (2025). A DeepAR-Based Modeling Framework for Probabilistic Mid-Long Term Streamflow Prediction. Water, 17, 2506.

Yan Y., Zhang W., Liu Y. and Li Z. (2023). Simulated annealing algorithm optimized GRU neural network for urban rainfall-inundation prediction. J. Hydroinformatics, 25, 1358–1379.

Zhang D., Quan J., Zhang H., Wang F., Wang H. and He X. (2015). Flash flood hazard mapping: A pilot case study in Xiapu River Basin, China. Water Sci. Eng., 8, 195–204.