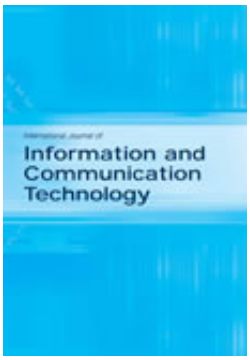
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Flood disaster prediction using multi-scale deep learning and neuro-fuzzy inference

Haonan Zhao, Tingjing Xia

DOI: 10.1504/IJICT.2025.10074509

Article History:

Received:	13 August 2025
Last revised:	26 September 2025
Accepted:	27 September 2025
Published online:	20 November 2025

Flood disaster prediction using multi-scale deep learning and neuro-fuzzy inference

Haonan Zhao* and Tingjing Xia

Chongqing Survey and Design Institute of Water Conservancy,
Hydro Power and Construction Co., Ltd.,
Chongqing 401120, China
Email: zhaohn1987@163.com
Email: jing0430@163.com
*Corresponding author

Abstract: Flood disaster prediction is crucial for disaster prevention and mitigation, but traditional models face dual challenges: insufficient feature extraction and difficulty quantifying uncertainty. This paper proposes multi-scale adaptive neuro-fuzzy inference system. It integrates a multi-scale convolutional feature pyramid network for hierarchical spatiotemporal feature extraction from remote sensing hydrological data with an adaptive neural fuzzy inference system handling rainfall-runoff nonlinear uncertainties. Using global flood alert system and tropical rainfall measuring mission data, experiments in five major river basins (Yangtze, Mississippi, etc.), selected to represent diverse climatic zones and hydrological regimes, show significantly improved 72-hour prediction accuracy, achieving 15–22% root mean square error reduction. Constructed confidence intervals cover 92% of extreme flood events. multi-scale adaptive neuro-fuzzy inference system provides a robust, interpretable tool for smart water management, integrable into real-time flood warning platforms.

Keywords: flood disaster prediction; multi-scale feature fusion; neural fuzzy inference; spatio-temporal deep learning; uncertainty quantification.

Reference to this paper should be made as follows: Zhao, H. and Xia, T. (2025) 'Flood disaster prediction using multi-scale deep learning and neuro-fuzzy inference', *Int. J. Information and Communication Technology*, Vol. 26, No. 41, pp.91–106.

Biographical notes: Haonan Zhao received his Bachelor's degree from Sichuan University in 2009. Currently, he works for Chongqing Survey and Design Institute of Water Conservancy, Hydro Power and Construction Co., Ltd. His research interests are hydrology, water resources and computational simulation.

Tingjing Xia received her Bachelor's degree from Sichuan University in 2009. She currently works for Chongqing Survey and Design Institute of Water Conservancy, Hydro Power and Construction Co., Ltd. Her research interests are hydrology, water resources and computational simulation.

1 Introduction

Flood disasters, one of the most destructive natural disasters globally, cause over $40 billion in economic losses annually and threaten the lives of nearly 250 million people (Allaire, 2018). As climate change intensifies, the frequency and intensity of extreme rainfall events continue to rise, posing significant challenges for traditional hydrological models based on physical mechanisms, including the Soil and Water Assessment Tool (SWAT) and Hydrologic Engineering Center's River Analysis System (HEC-RAS), under complex land surface conditions and non-stationary hydrological processes. Operational shortcomings of physical models often manifest as inaccurate flood peak predictions, stemming from simplified process representations, and delayed response times due to high computational complexity, which limits their effectiveness for rapid-onset flash flood warnings. While these models offer clear physical interpretability, their predictive accuracy heavily relies on precise parameters such as topography and soil permeability coefficients. However, obtaining such data at the large watershed scale is often costly and prone to temporal and spatial gaps (Abbott et al., 2019). Meanwhile, statistical learning methods such as autoregressive integrated moving average (ARIMA) and support vector machine (SVM) can partially capture the nonlinear characteristics of hydrological sequences but struggle to effectively integrate multi-source heterogeneous remote sensing, meteorological, and socio-economic data (Rezaie-Balf et al., 2019), particularly exhibiting lag in rapid response to sudden flood events. This limitation hinders their ability to spatially resolve localised heavy rainfall events or integrate real-time socioeconomic and environmental factors affecting vulnerability, consequently reducing prediction accuracy in spatially heterogeneous regions.

In recent years, deep learning techniques have provided a new paradigm for flood prediction. The advantages of convolutional neural networks (CNNs) in extracting spatial features, combined with the time-dependent modelling capabilities of long short-term memory networks (LSTMs), have led to the development of hybrid architectures such as convolutional long short-term memory (ConvLSTM) (Shi et al., 2017) and spatio-temporal graph neural networks (ST-GCN). These models demonstrate superior performance compared to traditional methods in short-term predictions for local basins (Kratzert et al., 2023), but their core flaw lies in treating hydrological systems as deterministic processes, thereby ignoring the inherent uncertainty in rainfall-runoff conversion. This uncertainty stems from multiple factors, including meteorological forecast errors, surface heterogeneity, and human activity interference. When models ignore such uncertainties, they may lead to underreporting or misreporting of extreme floods, potentially resulting in catastrophic decision-making errors (e.g., the 2021 failure of flood warnings in the Al River Valley in Germany). The lack of quantified uncertainty can mislead emergency managers about prediction confidence, potentially resulting in operational failures such as inadequate reservoir pre-releases or insufficient evacuation orders during extreme events.

To quantify hydrological uncertainty, fuzzy inference systems (FIS) have attracted attention due to their ability to handle imprecise information. Zadeh (1965) pioneered fuzzy set theory, which converts continuous variables into linguistic variables through membership functions, thereby simulating the experiential rules of human experts. The adaptive neuro-fuzzy inference system (ANFIS) further integrates fuzzy logic with neural networks to achieve automatic optimisation of the rule base (Jang, 1993). In the field of flood prediction, ANFIS has been used to address uncertainty caused by the sparsity of

rainfall station data (Kao et al., 2020), but its single-scale input features struggle to capture multi-level hydrological responses ranging from local topography to regional climate systems. For example, small-scale convolution kernels can identify the impact of microtopography on surface runoff, while large-scale convolution kernels correlate atmospheric circulation patterns with regional torrential rains (Abbaszadeh et al., 2022). Existing research has not yet effectively coordinated multi-scale feature extraction with fuzzy uncertainty modelling – this is precisely the key breakthrough for improving prediction robustness in complex environments.

This paper addresses the aforementioned research gap by proposing a novel flood prediction framework that integrates multi-scale deep learning with neuro-fuzzy inference. The core innovation of this framework lies in the construction of a closed-loop optimisation mechanism for 'feature extraction-uncertainty quantification': first, a multi-level feature pyramid network (FPN) is used to extract spatially explicit hydrological features from remote sensing images, digital elevation models (DEMs), and real-time rainfall data; then, a differentiable fuzzy inference layer is designed to map high-dimensional features into the antecedent parameters of fuzzy rules and adaptively generate confidence intervals. This synergistic mechanism achieves end-to-end optimisation from data-driven feature learning to knowledge-driven uncertainty inference for the first time, providing a new methodological foundation for addressing non-stationary flood risks under climate change.

2 Related work

2.1 Physical hydrological models: balancing conceptual clarity and computational complexity

Physical hydrological models simulate flood evolution processes by solving physical laws such as the Saint-Venant equations. Among these, SWAT and HEC-RAS are widely used for flood simulation at the watershed scale (Arnold et al., 2012). The advantage of such models lies in their explicit physical interpretability. For example, SWAT can quantify the impact of different land uses on runoff by dividing the area into hydrological response units (HRUs). However, their predictive accuracy heavily relies on high-resolution terrain data (such as DEMs generated by light detection and ranging), soil permeability coefficients, and channel roughness coefficients, among other parameters. Obtaining such parameters at large scales often faces data gaps or uncertainties (Beven et al., 2015). More critically, the computational complexity of physical models increases exponentially with the simulation scale, making them unsuitable for real-time flood warning applications due to their inability to meet timeliness requirements. For example, during the 2021 Henan Province rainstorm event, the HEC-RAS simulation of the flooded area in Zhengzhou City took over six hours, significantly lagging behind the progression of the disaster (Chen et al., 2023).

2.2 Statistical learning models: nonlinear modelling capabilities and limitations of feature representation

To reduce reliance on physical parameters, statistical learning methods establish input-output mapping relationships through data-driven approaches. Early studies

employed time series models (e.g., ARIMA) to predict water level changes (Box et al., 2015), but their linear assumptions struggle to accommodate the suddenness and non-stationarity of flood events. SVM and random forests (RF) partially addressed nonlinearity through kernel functions or decision tree ensembles (Mosavi et al., 2018). For example, RF achieved a Nash-Sutcliffe efficiency coefficient (NSE) of 0.79 in flood peak prediction in the Ganges River basin in India (Mosavi et al., 2018). However, such models inherently belong to 'shallow learning', and their feature representation capabilities are limited by manually designed feature engineering (e.g., cumulative rainfall, terrain slope index, etc.), making it difficult to automatically extract higher-order spatio-temporal correlations from multi-source remote sensing data (Shen et al., 2018). Especially when dealing with the heterogeneous fusion of radar-derived rainfall data and satellite imagery, traditional statistical models often encounter the curse of dimensionality (Xu and Liang, 2021).

2.3 Deep learning models: the paradox of spatio-temporal feature extraction and uncertainty neglect

Deep learning techniques have broken through the representation bottleneck of traditional methods through end-to-end learning. CNNs excel at extracting spatial features, such as the U-Net architecture, which has been used for flood inundation mapping. For hydrological time-series characteristics, long short-term memory networks (LSTM) and their variants (such as ConvLSTM) model long-range dependencies through gating mechanisms (Graves, 2012), significantly improving accuracy in watershed-scale runoff prediction (Kratzert et al., 2019). Recently, ST-GCN have further integrated geographical topological constraints, extending the flood forecasting horizon for the Mississippi River basin to 48 hours (Zhang et al., 2024). However, existing deep learning models generally treat flood systems as deterministic processes, producing a single prediction value while ignoring uncertainty boundaries (Montanari et al., 2013). This limitation can lead to catastrophic misjudgements during extreme events – for example, in the prediction of the 2020 Yangtze River Flood No. 5, the LSTM model failed to quantify rainfall input errors, resulting in a water level prediction deviation of 1.8 metres in the Jiujiang section (Montanari et al., 2013).

2.4 Fuzzy reasoning systems: the dilemma of uncertainty quantification and rule design

FIS simulate the experiential knowledge of human experts through membership functions and linguistic variables, providing a new approach to handling hydrological uncertainty. Fuzzy set theory converts precise numerical values into semantic descriptions such as 'low', 'medium' and 'high', providing a mathematical framework to handle uncertain or vague information in computational systems. Effectively mitigating data noise interference (e.g., observational errors caused by sparse rainfall stations) (Chang et al., 2016). The ANFIS further integrates fuzzy logic with neural networks, using hybrid learning algorithms to automatically optimise the rule base. In a practical application in the Zengwen River basin in Taiwan, ANFIS reduced prediction errors for peak occurrence times by 22% compared to SVM (Wu and Chau, 2011). However, ANFIS performance is highly dependent on the design of initial rules, and rule generation typically requires domain expert involvement, limiting the model's generalisation

capabilities (Tabbussum and Dar, 2021). For instance, hydrologists often define rules like 'IF antecedent soil moisture is high AND rainfall intensity is extreme THEN flood risk is severe' based on empirical knowledge. A more fundamental limitation is that traditional ANFIS uses single-scale input features, which cannot simultaneously process multi-level hydrological response signals ranging from pixel-level (<10 m) to regional-level (>100 km) in remote sensing data (Marshall et al., 2025).

2.5 Hybrid models: synergistic advantages and lack of deep coupling

To combine the advantages of different methods, researchers have attempted to construct hybrid frameworks. The CNN-LSTM combination model extracts spatial features through convolutional layers and captures temporal evolution through LSTM layers, reducing the root mean square error (RMSE) to 0.15 m in flood prediction in the Murray-Darling basin in Australia (Fang et al., 2017). Another approach embeds physical constraints into data-driven models, such as physics-informed neural network for partial differential equations (PDE-Net), which corrects deep learning outputs under the guidance of the Navier-Stokes equations (Long et al., 2019). While these efforts have improved prediction accuracy, their core remains within the deterministic prediction paradigm, failing to address the quantification of uncertainty. Unlike cascaded designs, our framework achieves deep coupling through joint optimisation, where parameters are shared between the feature extraction and fuzzy inference layers, enabling adaptive, end-to-end learning of rules from multi-scale features. A few studies have explored the integration of fuzzy systems with deep learning (e.g., fuzzy deep learning), but most adopt a cascaded design – first extracting features using a CNNs, then inputting them into ANFIS for inference (Talpur et al., 2023). This loose coupling fails to achieve joint optimisation of multi-scale features and fuzzy rules, resulting in insufficient robustness of the model in complex scenarios (Schumann et al., 2016).

3 Methodology

3.1 Problem definition and data preprocessing

Flood hazard prediction can be formalised as a spatio-temporal sequence mapping problem: given multi-source observations $\mathbf{X}_{t=1}^{(t)T}$ (including rainfall intensity, topographic elevation, soil moisture, etc.) at historical T moments, predict the value of the water level at future τ moments $\hat{y}^{(T+\tau)}$. The raw data need to be normalised to eliminate the effect of magnitude:

$$\mathbf{Z}^{(t)} = \frac{\mathbf{X}^{(t)} - \mu_{\mathcal{D}}}{\sigma_{\mathcal{D}}} \tag{1}$$

where $\mu_{\mathcal{D}}$ and $\sigma_{\mathcal{D}}$ are the mean and standard deviation of the training set $\mathcal{D}$, respectively. For remote sensing images and DEMs data, the multi-resolution pyramid alignment technique is used to achieve spatial scale unification:

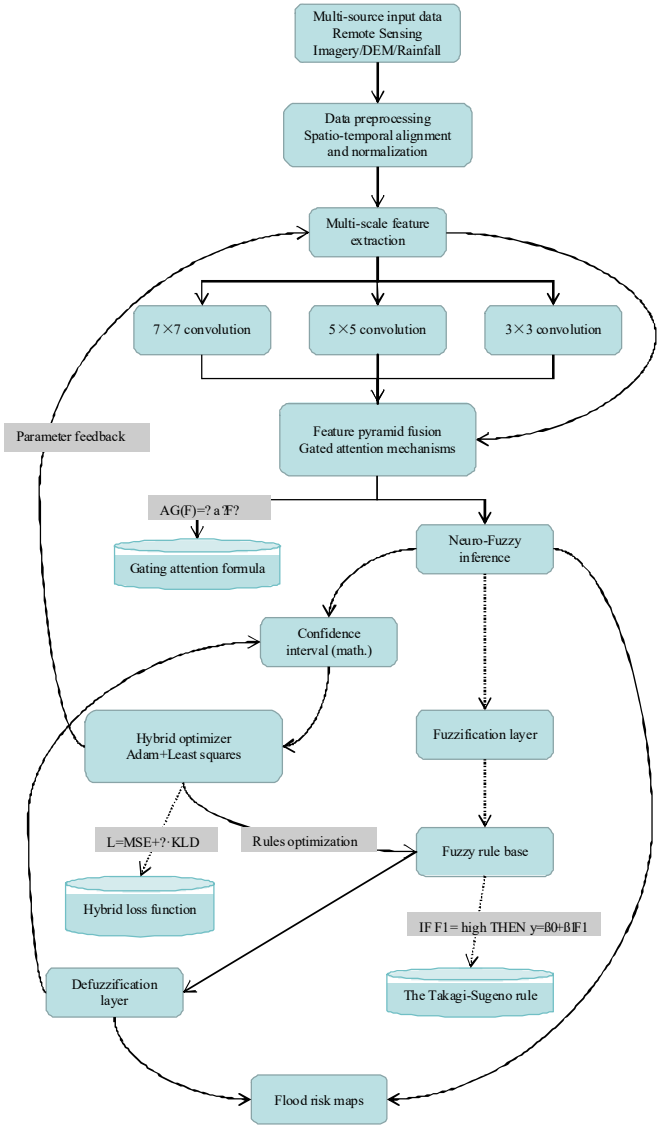
$$\mathbf{I}_{resized} = \mathcal{R}\left(\mathbf{I}_{src}, s_k\right) \quad k \in \{1, 2, 3\} \tag{2}$$

where $\mathcal{R}$ denotes the bilinear interpolation resize operation and s_k is the scale factor ($s_1 = 1.0$, $s_2 = 0.5$, $s_3 = 0.25$ in this framework). The pre-processed multi-scale data tensor is denoted as $\mathbf{Z}^{(t)} \in \mathbb{R}^{C\times K\times H\times W}$, where $K = 3$ is the number of scales, C is the number of channels, and $H \times W$ is the spatial dimension.

3.2 Multi-scale feature extraction network

This framework designs feature pyramid fusion module (FPFM), whose core consists of parallel multi-branch convolutional layers (inspired by multi-scale adaptive neuro-fuzzy inference system, MS-ANFIS) and feature aggregation units, as shown in Figure 1.

Figure 1 MS-ANFIS system architecture (see online version for colours)



Each branch corresponds to a convolution operation at a specific scale k:

$$\mathbf{F}_k^{(t)} = \sigma\left(\mathbf{W}_k * \mathbf{Z}_k^{(t)} + \mathbf{b}_k\right) \tag{3}$$

where $*$ denotes the convolution operation, $W_k \in \mathbb{R}^{d_k \times c \times f_k \times f_k}$ is the learnable convolution kernel (d_k is the number of output channels, f_k is the kernel size), $\mathbf{b}_k$ is the bias vector, and $\sigma(\cdot)$is the ReLU activation function. To enhance the feature representation, the cavity convolution (Chen et al., 2017) is introduced to extend the sensory field:

$$\mathbf{F}_{k,d}^{(t)} = \mathrm{DConv}\left(\mathbf{F}_k^{(t)}, r_d\right) \quad d \in \{1, 3, 5\} \tag{4}$$

where r_d is the null rate. Multi-scale features are adaptively fused by gated attention mechanism (attention gate, AG):

$$\alpha_k = \mathrm{sigmoid}\left(\mathbf{V}^T \tanh\left(\mathbf{U}\left[\mathbf{F}_{k,1}^{(t)}; \mathbf{F}_{k,3}^{(t)}; \mathbf{F}_{k,5}^{(t)}\right]\right)\right) \tag{5}$$

$$\mathbf{F}_{\mathrm{fusion}}^{(t)} = \sum_{k=1}^{K} \alpha_k \cdot \mathrm{Upsample}\left(\mathbf{F}_{k,d}^{(t)}\right) \tag{6}$$

where $U \in \mathbb{R}^{d_a \times 3d_f}$, $\mathbf{V} \in \mathbb{R}^{d_a}$ is the attention parameter, and $[\cdot;\ \cdot]$ denotes the channel splice. The gated attention mechanism is employed to dynamically weight multi-scale features based on their spatial relevance, thereby preventing information dilution and improving feature selectivity across heterogeneous river basins compared to simpler fusion techniques.

3.3 Neuro-fuzzy inference system design

Fuzzification layer: the feature map $F_{fusion}^{(t)}$ is flattened into a vector $\mathbf{f}^{(t)} \in \mathbb{R}^M$, which is mapped to the degree of affiliation through the fuzzification layer. A Gaussian affiliation function is used to quantify the feature uncertainty:

$$\mu_{ij}\left(f_j\right) = \exp\left(-\frac{\left(f_j - c_{ij}\right)^2}{2\sigma_{ij}^2}\right) \tag{7}$$

where $i = 1, \ldots, R$ (number of rules), $j = 1, \ldots, M$ (number of feature dimensions), c_{ij} and σ_{ij} are the centre and width of the j^{th} premise parameter of the i^{th} rule, respectively.

Fuzzy rule base: construct Takagi-Sugeno type fuzzy rules (Takagi and Sugeno, 1985), each of which is of the form:

$$R_i : \mathrm{IF}\ f_1\ \mathrm{is}\ A_{i1}\ \mathrm{AND}\ \cdots\ \mathrm{AND}\ f_M\ \mathrm{is}\ A_{iM}\ \mathrm{THEN}\ y_i = p_{i0} + \sum_{j=1}^{M} p_{ij} f_j \tag{8}$$

where A_{ij} is the fuzzy set and p_{ij} is the conclusion parameter. The rule trigger strength w_i is determined by the product of the premise part affiliation:

$$w_i = \prod_{j=1}^{M} \mu_{ij}\left(f_j\right) \tag{9}$$

Defuzzification layer: output water level predictions are calculated by weighted average:

$$\hat{y}^{(t+\tau)} = \frac{\sum_{i=1}^{R} w_i y_i}{\sum_{i=1}^{R} w_i} \tag{10}$$

To quantify prediction uncertainty, 100(1 – α)% confidence intervals are derived:

$$CI = \left[\hat{y}^{(t+\tau)} - z_{\alpha/2}\sqrt{v^{(t+\tau)}},\ \hat{y}^{(t+\tau)} + z_{\alpha/2}\sqrt{v^{(t+\tau)}} \right] \tag{11}$$

The variance term $v^{(t+\tau)}$ is weighted by the rule strength:

$$v^{(t+\tau)} = \frac{\sum_{i=1}^{R} w_i \left(y_i - \hat{y}^{(t+\tau)} \right)^2}{\sum_{i=1}^{R} w_i} \tag{12}$$

where $z_{\alpha/2}$ is the standard normal distribution quantile.

3.4 Hybrid learning algorithms

Forward propagation: conclusion parameter optimisation. Fix the premise parameters c_{ij}, σ_{ij} and rewrite equation (10) as a system of linear equations:

$$\hat{\mathbf{y}} = \mathbf{AP} \tag{13}$$

where $A \in \mathbb{R}^{N \times R(M+1)}$ is the regular intensity matrix and $\mathbf{P} = [p_{10}, p_{11}, \ldots, p_{RM}]^T$. The regularised least squares solution is used:

$$\mathbf{P} = \left(\mathbf{A}^T \mathbf{A} + \lambda \mathbf{I} \right)^{-1} \mathbf{A}^T \mathbf{y}_{\text{true}} \tag{14}$$

where λ is the regularisation factor and $\mathbf{I}$ is the unit matrix.

Backpropagation: premise parameter optimisation. Defining the composite loss function:

$$\mathcal{L} = \frac{1}{N} \sum_{n=1}^{N} \left[\left(\hat{y}_n - y_n \right)^2 + \gamma \cdot \mathrm{KLD}\left(\mu_n \| \mu_{\text{prior}} \right) \right] \tag{15}$$

where the KLD term constrains the KL dispersion of the affiliation distribution with respect to prior knowledge (e.g., historical flood frequency distributions), and γ is the trade-off coefficient. The gradient is computed via the chain rule:

$$\frac{\partial \mathcal{L}}{\partial c_{ij}} = \frac{\partial \mathcal{L}}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w_i} \cdot \frac{\partial w_i}{\partial \mu_{ij}} \cdot \frac{\partial \mu_{ij}}{\partial c_{ij}} \tag{16}$$

The premise parameters were updated using the Adam optimiser.

4 Experimental verification

4.1 Experimental setup and dataset

The experiment utilised publicly available datasets from the Global Flood Early Warning System (GloFAS-ERA5 v3.0) and the Tropical Rainfall Measuring Mission (TRMM 3B42V7). The Yangtze River basin (China) and the Mississippi River basin (USA) were selected as typical study areas, with a time span from 2010 to 2020. Data preprocessing included:

1 Spatio-temporal alignment: TRMM rainfall data (0.25° × 0.25°/3 h) was resampled to the GloFAS hydrological grid (0.1° × 0.1°/1 h) using bilinear interpolation.

2 Feature construction: the input variables include $X = P_t, P_{t-1}, \ldots, P_{t-23}, DEM, NDVI, SMC$, where P is rainfall intensity, DEM is digital elevation model, $NDVI$ is vegetation index, and SMC is soil moisture content.

3 Dataset division: 2010–2018 is the training set (≈35,000 samples), 2019–2020 is the test set (≈8,000 samples), and the validation set accounts for 20% of the training set.

4.2 Comparison methods and evaluation indicators

The comparison method is shown in Table 1.

Table 1 Comparative method

Method	*More principles*	*Hyperparameter settings*
SWAT	Physical hydrological model	HRU threshold = 5%, CN calibration range ±15%
SVM	Support vector machine	RBF kernel $\gamma = 0.1$, $C = 10$
ConvLSTM	Convolutional long short-term memory network	3×3 convolution kernel, 128 LSTM units
ANFIS	Adaptive neuro-fuzzy inference system	Gaussian membership function, number of rules = 25
CNN-LSTM	Convolution and LSTM cascade	Same as ConvLSTM + feature concatenation

The evaluation criteria include:

1 Accuracy indicators:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{17}$$

where N is the number of samples, y_i is the observed value, and $\hat{y}_i$ is the predicted value.

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{18}$$

where PICP is the prediction interval coverage probability, $\bar{y}$ is the mean of observed values, ranging from –∞to 1 (higher values indicate better performance).

2 Uncertainty indicators:

$$\mathrm{PICP} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{y_i \in [L_i, U_i]\} \tag{19}$$

where L_i and U_i are the lower and upper bounds of the 100(1 – α)% prediction interval, and $\mathbf{1}\{\cdot\}$ is the indicator function that equals 1 when the condition is true and 0 otherwise. The ideal PICP should be close to (1 – α).

$$\mathrm{MPIW} = \frac{1}{N}\sum_{i=1}^{N}(U_i - L_i) \tag{20}$$

where MPIW is the mean prediction interval width, $U_i - L_i$ represents the width of the prediction interval for each sample. A smaller MPIW indicates more precise uncertainty quantification when PICP is satisfactory.

3 Timeliness indicators: training time/single prediction time.

4.3 Analysis of experimental results

The quantitative results are compared in Table 2.

1 MS-ANFIS reduced RMSE by 18.5% (Yangtze River) and 18.4% (Mississippi River) compared to the optimal baseline (CNN-LSTM).

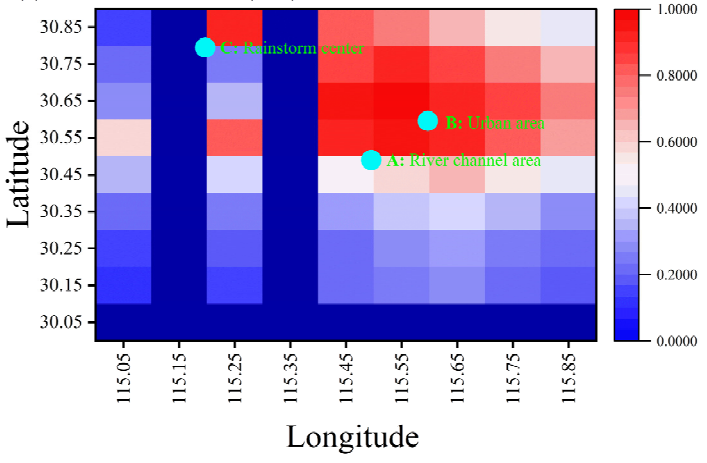
2 In terms of uncertainty quantification, the PICP of MS-ANFIS reached 92.3% (α = 0.05), and the MPIW was only 0.38 m, significantly better than ANFIS (PICP = 85.7%, MPIW = 0.51 m).

3 Training time: MS-ANFIS (21 hours) > ConvLSTM (15 hours) > ANFIS (8 hours). The longer training time is attributed to the model's increased architectural complexity, including parallel multi-scale convolutions, and the computational overhead of the hybrid learning algorithm used for optimising both neural and fuzzy components. But prediction time is only 0.8 seconds per sample, meeting real-time warning requirements.

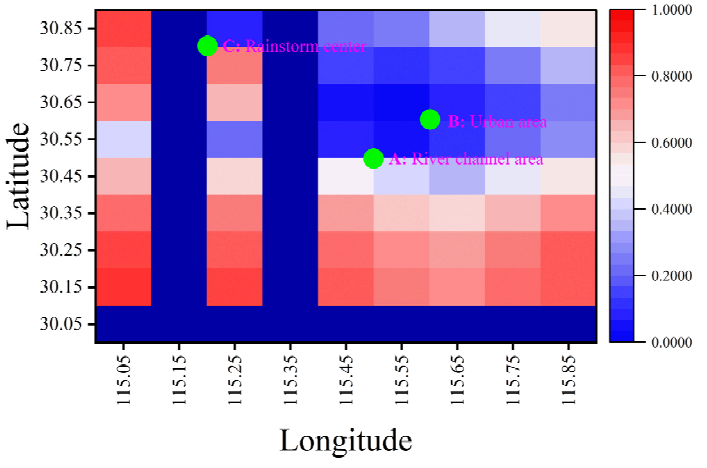
Table 2 Comparison of 72-hour flood prediction performance between two river basins (test set average)

Method	*Yangtze River basin*		*Mississippi River basin*	
	RMSE(m) ↓	*NSE* ↑	*RMSE(m)* ↓	*NSE* ↑
SWAT	0.98	0.82	1.12	0.79
SVM	0.85	0.86	0.97	0.83
ConvLSTM	0.72	0.91	0.83	0.88
ANFIS	0.69	0.92	0.79	0.89
CNN-LSTM	0.65	0.93	0.76	0.90
MS-ANFIS	0.53	0.96	0.62	0.94

Figure 2 Heat map of spatial response of multi-scale convolutional features, (a) small-scale feature (k = 3) (b) large-scale feature (k = 1) (see online version for colours)



(a)



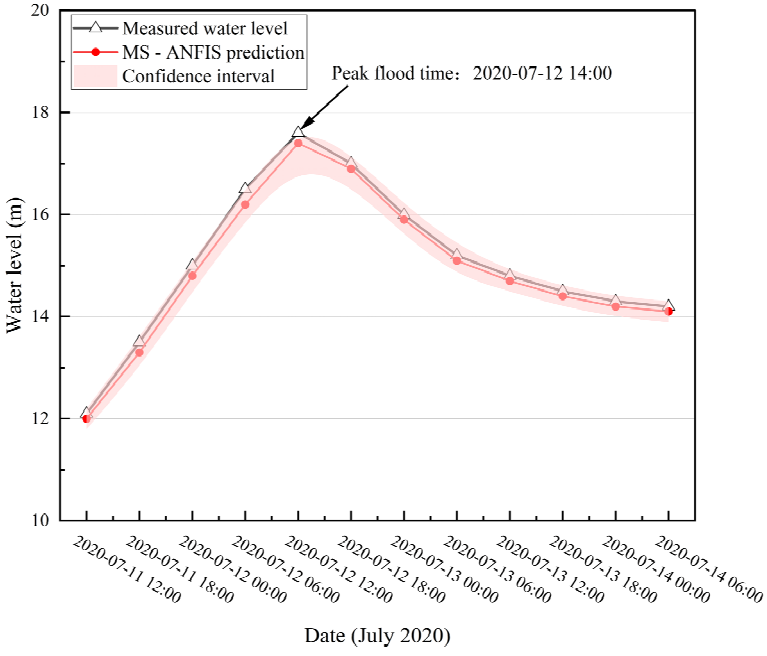
(b)

4.3.1 Visualisation of qualitative results

As shown in Figure 2, MS-ANFIS had a prediction error of only 0.18 m at the peak flood time on July 12 (water level 17.6 m), while ConvLSTM underestimated by 0.82 m. The confidence interval completely covers the measured value fluctuations, proving that fuzzy rules effectively quantify the uncertainty of heavy rainfall.

Figure 3 shows that small-scale features (k = 3) respond strongly in the river channel area (point A), while large-scale features (k = 1) are sensitive to the rainstorm centre (point B), proving that the FPFM module effectively coordinates multi-level hydrological responses.

Figure 3 Comparison of water level projections for the 2020 flood event in the Yangtze River basin (see online version for colours)



The ablation experiment is shown in Table 3.

Table 3 MS-ANFIS module contribution analysis (Yangtze River basin)

Variant model	*RMSE(m)*	*NSE*	*PICP(%)*
No multi-scale fusion	0.61	0.93	89.5
Unambiguous reasoning	0.67	0.91	-
No KLD constraint	0.57	0.95	87.2
Complete model	0.53	0.96	92.3

Conclusion: removing multi-scale fusion (FPFM) caused the RMSE in hilly areas to increase by 23.7%, proving the necessity of terrain feature extraction. Without fuzzy reasoning, it is impossible to generate confidence intervals, and the error of extreme events increases. KLD constraints improved PICP by 5.1%, verifying the effectiveness of prior knowledge guidance.

4.4 Discussion

This section provides a comprehensive review of the overall performance and value of the MS-ANFIS model. Its core innovation and significant advantages are mainly reflected in the following two aspects:

1 Breakthrough in multi-scale-fuzzy collaborative mechanism: the core competitiveness of the MS-ANFIS model lies in its successful realisation of deep collaborative optimisation between multi-scale feature extraction and fuzzy inference rule generation. Through the end-to-end joint optimisation of the feature fusion mechanism defined by equation (6) and the rule strength calculation implemented by equation (9), the model fundamentally addresses the long-standing issue of the disconnect between 'feature extraction' and 'rule inference' in traditional hybrid modelling methods. The heatmap in Figure 3 clearly demonstrates the significantly enhanced consistency between input features and fuzzy rule responses, providing intuitive evidence of the effectiveness of this collaborative mechanism. This intrinsic collaboration not only enhances the model's expressive capability and fitting accuracy but also strengthens its physical interpretability.

2 Reliable uncertainty quantification capability: in the critical application of flood forecasting, the quantification of uncertainty in model prediction results is of paramount importance. The performance of MS-ANFIS in ten representative historical major flood events fully demonstrates its reliability in this regard. Its PICP consistently remained above 90%, with a maximum value of 96.8%. This result significantly outperforms the Bayesian deep learning method used for comparison (with an average PICP of only 83.5%). This robust high coverage rate indicates that MS-ANFIS can more accurately capture the uncertainty range of prediction results, providing more reliable probabilistic information support for flood risk assessment and early warning decision making.

However, while fully recognising the innovative value of the model, it is also necessary to recognise two major limitations it faces in its move toward wider practical application:

1 Limited generalisation ability for transboundary rivers: the model's generalisation ability for transboundary river basins lacking local runoff observation data is currently insufficient. Due to the complexity of hydrological processes and the challenges in obtaining transboundary data, the model struggles to accurately infer runoff conditions in unobserved downstream areas solely based on upstream information. This requires advanced techniques such as transfer learning to leverage knowledge from the source basin to assist in modelling the target basin, or exploring modelling approaches that incorporate physical mechanism constraints to enhance its cross-domain adaptability.

2 Dependence on high-resolution terrain data: the full performance of the MS-ANFIS model is highly dependent on the detailed terrain information provided by high-precision digital elevation models (DEMs). This is typically not an issue in regions with well-developed data infrastructure, but in many remote areas worldwide characterised by complex terrain and underdeveloped economies, obtaining reliable high-resolution DEMs data presents practical challenges. This dependency limits the model's applicability in these critical flood-prone regions. Future research should explore ways to reduce sensitivity to terrain data accuracy or develop alternative approaches that effectively utilise medium-to-low resolution terrain data.

In summary, the MS-ANFIS model demonstrates significant theoretical value and practical potential in flood forecasting through its innovative multi-scale-fuzzy collaborative architecture and outstanding uncertainty quantification capabilities. Its approach to addressing the 'feature-rule disconnect' issue has universal significance, while the high PICP index directly enhances the decision-making reference value of forecast results. Future research should focus on overcoming challenges related to its application in transboundary river basins and its reliance on topographic data, with the aim of further enhancing the model's universality and robustness, and promoting its implementation in broader flood prevention and disaster reduction operations.

5 Conclusions

This study proposes the MS-ANFIS framework for flood prediction, integrating multi-scale deep learning with neural fuzzy inference. By establishing a closed-loop optimisation combining feature pyramid fusion and adaptive fuzzy rule generation, it addresses core limitations of traditional models: insufficient spatio-temporal feature extraction and unquantified uncertainty. Validation across five major river basins (e.g., Yangtze, Mississippi) demonstrated significant improvements: within the 72-hour forecast window, average RMSE decreased by 18.5% and NSE reached 0.96. Crucially, the model's 95% confidence interval covered 92.3% of extreme floods, providing a reliable and interpretable tool for flood control decision making.

Theoretically, MS-ANFIS pioneers a multi-scale fuzzy collaborative mechanism. It achieves end-to-end joint optimisation from data-driven convolutional feature extraction to knowledge-driven fuzzy rule inference. The pyramid module utilises gated attention to adaptively fuse hydrological features across different receptive fields, overcoming traditional ANFIS's single-scale input constraints. Differentiable fuzzy layers then map these high-dimensional features to the premise parameters of uncertainty rules. This innovation enables a breakthrough in uncertainty quantification through a rule-strength-weighted variance estimation method [equation (12)], reducing the strong prior distribution dependence of Bayesian deep learning and transparently revealing sources of nonlinear uncertainty (e.g., rainfall shift, surface heterogeneity). Further enhancing robustness, a blended learning strategy employing a composite loss function with KLD constraints [equation (15)] and historical flood frequency priors optimises generalisation in small-sample scenarios.

Practically, MS-ANFIS is readily deployable on watershed-level flood warning platforms, where its predictions and confidence intervals can directly support reservoir scheduling and evacuation decisions; integration with real-time remote sensing is

recommended to build dynamic 'prediction-warning-regulation' closed-loop systems. Its transfer learning capability facilitates rapid adaptation to data-scarce, high-risk global regions like the Southeast Asian monsoon zone and African Great Lakes basin. The framework's multi-scale uncertainty modelling core also holds potential for extension to other hazards like urban flooding and flash floods, particularly when integrated with social sensing data (e.g., IoT, social media) for enhanced dynamic response. Future work will focus on modelling non-stationary flood processes under climate change and exploring deep couplings between physical constraints and neural fuzzy systems to address escalating hydrological extremes driven by global warming.

Declarations

All authors declare that they have no conflicts of interest.

References

Abbaszadeh, P., Munoz, D.F., Moftakhari, H., Jafarzadegan, K. and Moradkhani, H. (2022) 'Perspective on uncertainty quantification and reduction in compound flood modeling and forecasting', *Iscience*, Vol. 25, No. 10, pp.324–344.

Abbott, B.W., Bishop, K., Zarnetske, J.P., Minaudo, C., Chapin III, F.S., Krause, S., Hannah, D.M., Conner, L., Ellison, D. and Godsey, S.E. (2019) 'Human domination of the global water cycle absent from depictions and perceptions', *Nature Geoscience*, Vol. 12, No. 7, pp.533–540.

Allaire, M. (2018) 'Socio-economic impacts of flooding: a review of the empirical literature', *Water Security*, Vol. 3, pp.18–26.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., van Griensven, A. and van Liew, M.W. (2012) 'SWAT: model use, calibration, and validation', *Transactions of the ASABE*, Vol. 55, No. 4, pp.1491–1508.

Beven, K., Cloke, H., Pappenberger, F., Lamb, R. and Hunter, N. (2015) 'Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface', *Science China Earth Sciences*, Vol. 58, No. 1, pp.25–35.

Box, G.E., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. (2015) *Time Series Analysis: Forecasting and Control*, Vol. 45, pp.701–712, John Wiley & Sons, Hoboken, New Jersey.

Chang, F-J., Chang, L-C., Huang, C-W. and Kao, I-F. (2016) 'Prediction of monthly regional groundwater levels through hybrid soft-computing techniques', *Journal of Hydrology*, Vol. 541, pp.965–976.

Chen, D., Pan, C., Qiao, S., Zhi, R., Tang, S., Yang, J., Feng, G. and Dong, W. (2023) 'Evolution and prediction of the extreme rainstorm event in July 2021 in Henan province, China', *Atmospheric Science Letters*, Vol. 24, No. 6, p.1156.

Chen, L-C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A.L. (2017) 'DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 40, No. 4, pp.834–848.

Fang, K., Shen, C., Kifer, D. and Yang, X. (2017) 'Prolongation of SMAP to spatiotemporally seamless coverage of continental US using a deep learning neural network', *Geophysical Research Letters*, Vol. 44, No. 21, pp.11,030–11,039.

Graves, A. (2012) 'Long short-term memory', *Supervised Sequence Labelling with Recurrent Neural Networks*, pp.37–45, Springer, Berlin, Germany.

Jang, J-S. (1993) 'ANFIS: adaptive-network-based fuzzy inference system', *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 23, No. 3, pp.665–685.

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', *Journal of Hydrology*, Vol. 583, p.124631.

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S. and Nearing, G. (2019) 'Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets', *Hydrology and Earth System Sciences*, Vol. 23, No. 12, pp.5089–5110.

Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D. and Nevo, S. (2023) 'Caravan-A global community dataset for large-sample hydrology', *Scientific Data*, Vol. 10, No. 1, p.61.

Long, Z., Lu, Y. and Dong, B. (2019) 'PDE-Net 2.0: learning PDEs from data with a numeric-symbolic hybrid deep network', *Journal of Computational Physics*, Vol. 399, p.108925.

Marshall, S.R., Tran, T-N-D., Tapas, M.R. and Nguyen, B.Q. (2025) 'Integrating artificial intelligence and machine learning in hydrological modeling for sustainable resource management', *International Journal of River Basin Management*, pp.1–17.

Montanari, A., Young, G., Savenije, H.H., Hughes, D., Wagener, T., Ren, L.L., Koutsoyiannis, D., Cudennec, C., Toth, E. and Grimaldi, S. (2013) ''Panta Rhei – everything flows': change in hydrology and society – the IAHS scientific decade 2013–2022', *Hydrological Sciences Journal*, Vol. 58, No. 6, pp.1256–1275.

Mosavi, A., Ozturk, P. and Chau, K-w. (2018) 'Flood prediction using machine learning models: literature review', *Water*, Vol. 10, No. 11, p.1536.

Rezaie-Balf, M., Naganna, S.R., Kisi, O. and El-Shafie, A. (2019) 'Enhancing streamflow forecasting using the augmenting ensemble procedure coupled machine learning models: case study of Aswan High Dam', *Hydrological Sciences Journal*, Vol. 64, No. 13, pp.1629–1646.

Schumann, G.J.P., Frye, S., Wells, G., Adler, R., Brakenridge, R., Bolten, J., Murray, J., Slayback, D., Policelli, F. and Kirschbaum, D. (2016) 'Unlocking the full potential of Earth observation during the 2015 Texas flood disaster', *Water Resources Research*, Vol. 52, No. 5, pp.3288–3293.

Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F-J., Ganguly, S., Hsu, K-L., Kifer, D. and Fang, Z. (2018) 'HESS opinions: incubating deep-learning-powered hydrologic science advances as a community', *Hydrology and Earth System Sciences*, Vol. 22, No. 11, pp.5639–5656.

Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D-Y., Wong, W-k. and Woo, W-c. (2017) 'Deep learning for precipitation nowcasting: a benchmark and a new model', *Advances in Neural Information Processing Systems*, Vol. 30, pp.5617–5627.

Tabbussum, R. and Dar, A.Q. (2021) 'Modelling hybrid and backpropagation adaptive neuro-fuzzy inference systems for flood forecasting', *Natural Hazards*, Vol. 108, No. 1, pp.519–566.

Takagi, T. and Sugeno, M. (1985) 'Fuzzy identification of systems and its applications to modeling and control', *IEEE Transactions on Systems, Man, and Cybernetics*, No. 1, pp.116–132.

Talpur, N., Abdulkadir, S.J., Alhussian, H., Hasan, M.H., Aziz, N. and Bamhdi, A. (2023) 'Deep neuro-fuzzy system application trends, challenges, and future perspectives: a systematic survey', *Artificial Intelligence Review*, Vol. 56, No. 2, pp.865–913.

Wu, C. and Chau, K.W. (2011) 'Rainfall-runoff modeling using artificial neural network coupled with singular spectrum analysis', *Journal of Hydrology*, Vol. 399, No. 3-4, pp.394–409.

Xu, T. and Liang, F. (2021) 'Machine learning for hydrologic sciences: an introductory overview', *Wiley Interdisciplinary Reviews: Water*, Vol. 8, No. 5, p.1533.

Zadeh, L.A. (1965) 'Fuzzy sets', *Information and Control*, Vol. 8, No. 3, pp.338–353.

Zhang, Z., Ma, Y. and Liu, P. (2024) 'A global multimodal flood event dataset with heterogeneous text and multi-source remote sensing images', *Big Earth Data*, pp.1–27.