https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

The impact of spatial resolution on hourly flood modeling in large watersheds

Lei Ye[1], Xiaoyang Li[1, *], Jilie Li[1], Chi Zhang[1], Huicheng Zhou[1]

[1] School of Infrastructure Engineering, Dalian University of Technology, Dalian 116024, China

Correspondence to: Xiaoyang Li (lixiaoyang1998@mail.dlut.edu.cn)

Abstract.

The spatial resolution of hydrological modeling is a critical factor affecting flood simulation accuracy, especially in large watersheds characterized by complex watershed characteristics. However, its influence on the accuracy of hourly flood simulations at both watershed outlets and internal locations remains insufficiently understood, hindering rational spatial-resolution selection for large-scale flood forecasting. This study evaluates hourly flood simulations across five spatial resolutions (1 km, 3 km, 5 km, 10 km, and sub-watershed) at the watershed outlet and multiple internal stations in the Jialing River Basin, China (157,000 km²). An XGBoost-based model is employed to identify flood characteristics sensitive to spatial resolution and to quantify their nonlinear effects on simulation accuracy. Based on these relationships, spatial-resolution recommendations are derived for different flood-characteristic categories, and the effectiveness of spatial refinement under coarse rainfall inputs is examined. Results show that spatial refinement markedly improves simulation accuracy at internal locations but yields only marginal gains at the watershed outlet. Watershed area is identified as the dominant factor governing resolution sensitivity, while rainfall characteristics and underlying-surface properties exert strong nonlinear influences. Fine grids (1–3 km) are most effective under flood conditions with strong nonlinearity, but their advantages diminish rapidly as rainfall inputs become coarser, indicating that increased spatial resolution cannot compensate for insufficient rainfall information. Overall, these findings advance current understanding of spatial-resolution effects on hourly flood simulations and provide practical guidance for spatial-resolution selection in large-watershed modeling.

1 Introduction

Large watersheds cover vast areas with highly heterogeneous rainfall and complex watershed characteristics, often giving rise to localized flooding (Fischer & Schumann, 2021). These localized floods converge from tributaries into the main river and can ultimately develop into major flood events at the watershed outlet. Consequently, flood forecasting in large watersheds requires modeling flood processes both at the outlet and within the watershed (Ma et al., 2024; Fischer & Schumann, 2021). In commonly used distributed hydrological models, large watersheds are discretized into spatial resolutions of 10 km, 5 km, or even 1 km to simulate flood processes at the grid cell (Zhu et al., 2024). Finer spatial resolutions can better represent runoff generation and flow routing, particularly for localized flood responses. However, research shows that finer resolutions do not always improve simulation accuracy, as their effectiveness is limited by the

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

spatial resolution of input data (Jiang et al., 2025; Barnhart et al., 2024; Tian et al., 2020). Therefore, accurately assessing the influence of spatial resolution on flood simulations and selecting an appropriate resolution remain key challenges.

Existing studies on spatial-resolution effects can be broadly divided into two types: those focusing on flood simulations at the watershed outlet, and those that also examine localized flooding within the watershed (Barnhart et al., 2024; Lobligeois et al., 2014). Outlet-focused studies indicate that refining spatial resolution does not significantly improve daily flood simulations (Mateo et al., 2017; Zhu et al., 2024). In contrast, localized floods are more strongly influenced by spatial heterogeneity of rainfall and the underlying surface, requiring finer spatial discretization (Douinot et al., 2016), with greater benefits of fine resolution consistently observed in smaller watersheds (Jiang et al., 2025; Qiao et al., 2019). Notably, most existing studies on spatial scale effects focus on daily floods (Tian et al., 2020; Mateo et al., 2017). However, localized floods in large watersheds propagate rapidly and require hourly modeling, while daily simulations smooth rainfall variability and flood peaks and cannot be directly used for spatial-resolution selection at the hourly scale (Zhu et al., 2024; Li et al., 2024). This gap motivates a systematic evaluation of spatial-resolution effects on hourly flood simulations at both watershed outlets and internal locations.

The selection of spatial resolution depends on differences in flood simulation accuracy across multiple spatial resolutions. In practice, high-resolution modeling is warranted only when spatial refinement leads to a substantial improvement in simulation performance (Barnhart et al., 2024; Lobligeois et al., 2014). For example, in upstream drainage areas smaller than 100 km², a 0.05° (~5 km) resolution yields higher daily simulation accuracy (KGE) than coarser resolutions of 0.083°, 0.1°, and 0.25° (Jiang et al., 2025). Similarly, when rainfall exhibits strong spatiotemporal heterogeneity, finer spatial resolutions (e.g., 1 km) outperform coarser grids (e.g., 3 km and 9 km) in simulating flood processes (Tian et al., 2020; Yu et al., 2014). In addition, the benefits of finer spatial resolutions depend strongly on underlying surface characteristics, particularly under dry soil-moisture and high-elevation conditions (Aerts et al., 2022; Shrestha et al., 2015). However, existing studies on spatial refinement typically emphasize individual rainfall or surface indicators, without clarifying the importance of different flood characteristics in spatial-resolution selection. Therefore, identifying the key characteristic governing the choice of spatial modeling resolution is essential.

Beyond flood characteristics, flood simulation accuracy is also strongly constrained by the spatial resolution of modeling input data. Commonly used surface datasets, such as SRTM elevation data (30 m), GlobeLand30 land-cover data (30 m), and HWSD soil data (1 km), are sufficient to support fine-resolution hydrological modeling (Simard et al., 2024; Lovat et al., 2019). In contrast, rainfall is the most critical input for hydrological models, yet its spatial resolution is often insufficient to meet the requirements of high-resolution (e.g., 1 km) modeling (Towner et al., 2019; Fraga et al., 2019). Numerous studies show that coarse rainfall inputs introduce substantial uncertainty and reduce flood simulation accuracy, particularly in smaller watersheds (Huang et al., 2019; Michelon et al., 2021). However, it remains unclear whether refining spatial resolution can still provide relative gains in simulation accuracy under sparse rainfall station density conditions. Hence, understanding how rainfall-station density modulates flood simulation accuracy across spatial resolutions is essential for selecting appropriate modeling resolutions.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

This study focuses on the Jialing River Basin in China (157,000 km^2) and evaluates hourly flood simulations at both the watershed outlet and internal locations using five spatial resolutions (1 km, 3 km, 5 km, 10 km, and sub-watershed). An XGBoost model is employed to identify flood characteristics sensitive to spatial resolution and to examine how simulation accuracy varies with these factors across different resolutions. In addition, the effects of reduced rainfall station density (25%, 50%, and 75% reductions) on flood simulation accuracy are systematically investigated across spatial resolutions. This study aims to address three key questions:

1. How does modeling spatial resolution affect hourly flood simulation accuracy at the watershed outlet and internal locations?
2. Which factors are critical in selecting appropriate spatial resolutions, and how do they influence flood simulation accuracy?
3. Under decreasing rainfall station density, can refining spatial resolution still yield relative gains in flood simulation accuracy?

The remainder of this paper is organized as follows: Section 2 describes the data and methods, including the hydrological model and the experimental setup for evaluating spatial resolution effects. Section 3 presents the hydrological modeling results across different spatial resolutions, while Section 4 provides a detailed analysis and discussion of the results. Finally, Section 5 summarizes the main conclusions and outlines the study's limitations.

2 Data and Methods

2.1 Study area and data

The Jialing River, a major left-bank tributary of the upper Yangtze River, drains an area of approximately 157,000 km², spanning 102°30′–109°E and 29°40′–34°30′N (Fig. 1). The watershed lies in a subtropical monsoon climate zone where frequent heavy summer rainfall increases both the likelihood and intensity of floods. The region is characterized by hilly and mountainous terrain, pronounced topographic relief, numerous tributaries, and steep river gradients, all of which accelerate runoff and heighten susceptibility to localized flooding. Fig. 1 shows the spatial distribution of 321 rainfall stations. To ensure representative flood data, ten streamflow stations were selected across a wide range of control areas (233–157,014 km²). The Beibei (BB) station at the watershed outlet records an average annual discharge of 2,120 m³/s and peak flows up to 44,800 m³/s. Tingzikou (TZK) is the only control reservoir on the main stem, while Luoduxi (LDX) and Xiaoheba (XHB) are located downstream of major tributaries. The Bazhong (BZ), Maoba (MB), Hufeng (HF), Yongchang (YC), Zhaojiaci (ZJC), and Jiange (JG) stations lie in small to medium watersheds (233–2,707 km²) and are particularly prone to flooding triggered by localized intense rainfall.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

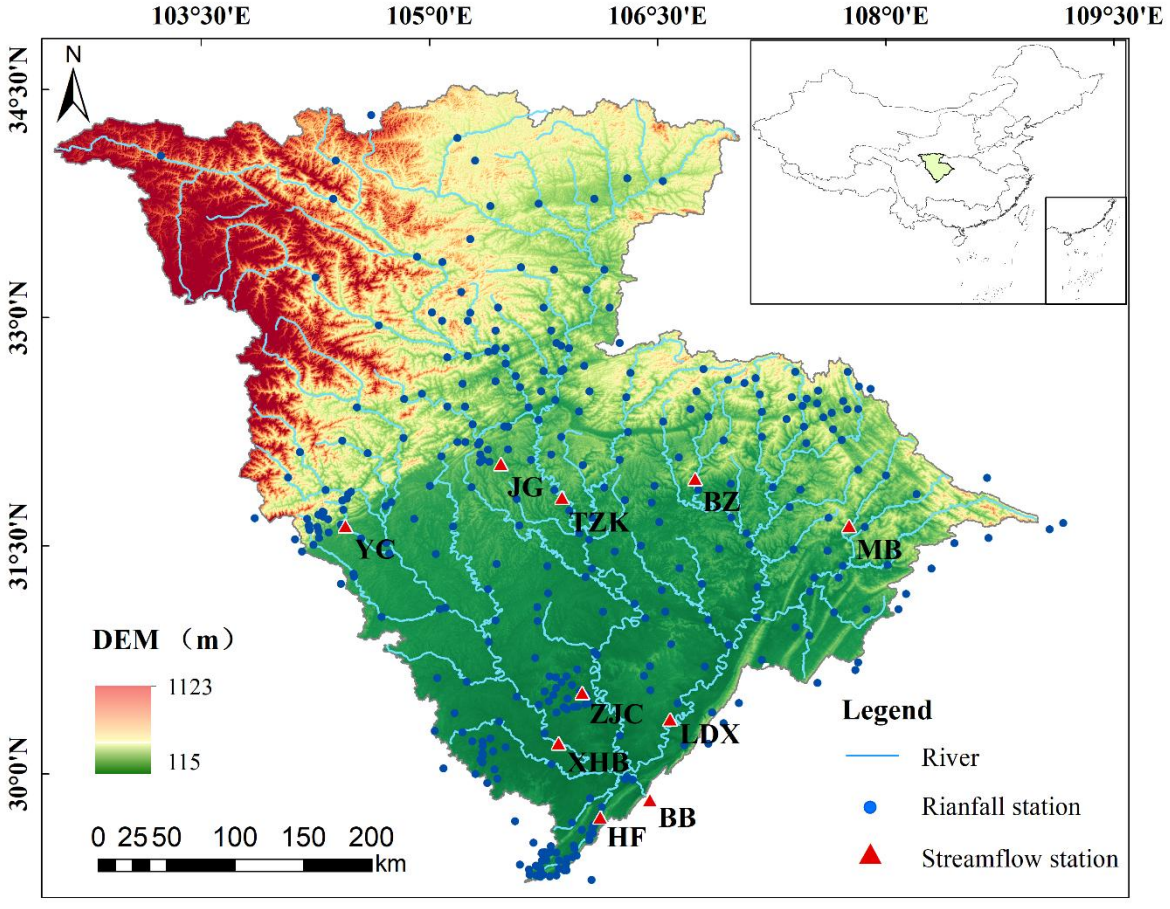


Figure 1. Location of the study area and the distribution of rainfall and streamflow stations.

Hydrological modeling requires multiple types of input data, including rainfall, evaporation, and underlying surface characteristics. Hourly rainfall and streamflow data from the flood season (May–October) during 2008–2021 were used, as summarized in Table 1. Because some stations have shorter observation periods, their records begin after 2008. Rainfall station density across the hydrological stations ranges from 29.1 to 961.4 km2·gauge-1. TZK, located in sparsely populated mountainous terrain, has the lowest density, whereas downstream regions with dense populations and frequent localized floods exhibit much higher densities. Overall, the rainfall network provides adequate spatial coverage across the watershed.

A total of 282 flood events were identified based on the temporal correlation of storm–flood processes among stations. A 30 m Digital Elevation Model (DEM) was obtained from the United States Geological Survey (USGS). The 1 km soil dataset was sourced from the Food and Agriculture Organization (FAO). Evaporation data were taken from the 1 km monthly potential evapotranspiration dataset for China, which was derived using the Hargreaves formula and 1 km temperature data (Peng et al., 2019). Information on the locations and storage capacities of small and medium-sized reservoirs was obtained from the China Reservoir Dataset (Song et al., 2022).

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Table 1. Rainfall and flood data statistics at hydrological stations in the study watersheds

Hydrological station	Abbreviation	Control area (km^2)	Number of rainfall stations	rainfall station density ($km^2 \cdot gauge^{-1}$)	Time	Flood events
Beibei	BB	157041	321	489.2	2008-2021	44
Tingzikou	TZK	62490	65	961.4	2013-2021	12
Luoduxi	LDX	37682	80	471.0	2008-2021	51
Xiaoheba	XHB	28826	48	600.5	2008-2021	28
Bazhong	BZ	2707	11	246.1	2008-2021	54
Maoba	MB	1415	13	108.8	2008-2021	42
Hufeng	HF	762	10	76.2	2014-2021	13
Yongchang	YC	539	9	59.9	2014-2021	14
Zhaojiaci	ZJC	441	8	55.1	2013-2021	10
Jiange	JG	233	8	29.1	2010-2021	14

2.2 Grid-based hydrological model

Choosing an appropriate hydrological model depends on the runoff-generation and routing characteristics of the study area. The study watershed spans multiple climate zones, and the highly uneven spatiotemporal distribution of rainfall leads to diverse and complex runoff-generation mechanisms. The routing process also exhibits strong nonlinear behavior due to the predominance of mountainous terrain and substantial topographic variability. In addition, numerous hydraulic projects within the basin further influence flood dynamics. Considering these characteristics, this study employs the gridded Dahuofang (GDHF) hydrological model, which incorporates the spatial distribution of rainfall, soil properties, and reservoir regulation, making it well-suited for simulating hourly flood events in the study area (Li et al., 2024).

Fig. 2 illustrates the soil delineation and runoff generation process of the GDHF model. Soil is divided into three layers—surface, lower, and deep—while a two-layer infiltration curve represents the spatial heterogeneity of the total infiltration rate f and the lower-layer infiltration rate f_u, as shown in Figs. 2(b) and 2(c). These infiltration characteristics are expressed as:

$$\alpha_1 = 1 - (1 - f/F_1)^{B-1} \quad (1)$$

$$\alpha_2 = 1 - (1 - f_u/F_2)^{B-1} \quad (2)$$

where α_1 and α_2 denote the fractions of the area with infiltration rates below f and f_u, respectively; F_1 and F_2 are the maximum total and lower-layer infiltration rates; and B is the shape parameter.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Hillslope routing is simulated using a two-parameter Gamma-distributed unit hydrograph. To better represent the nonlinear behavior of channel flow, a time-varying distributed unit hydrograph that incorporates rainfall intensity is used. Detailed descriptions of the GDHF runoff-generation and routing algorithms can be found in Section 2.2 of Li et al. (2024

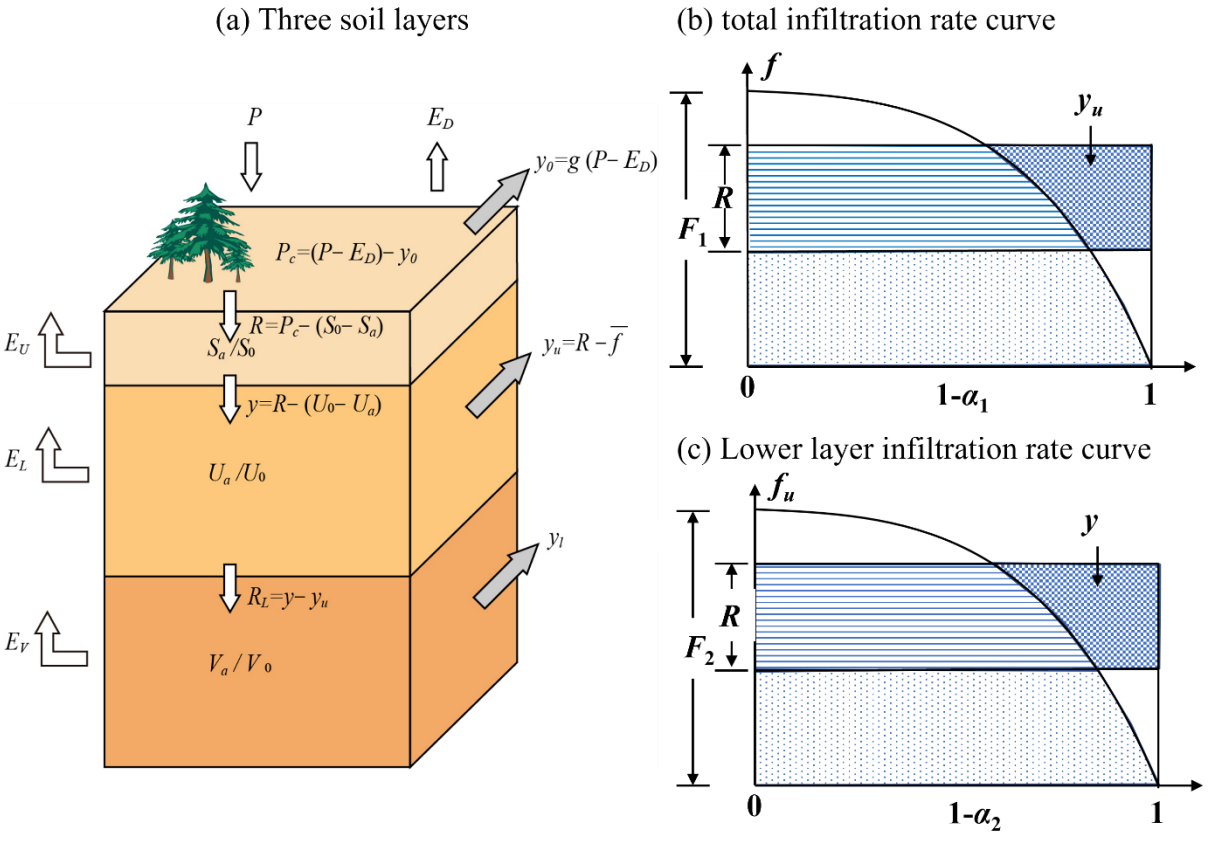


Figure 2. The structural diagram of the GDHF hydrological model.

2.3 Experimental design for comparing hourly flood simulations under different spatial resolutions

To evaluate the impact of spatial resolution on hourly flood simulations, the experimental design is shown in Fig. 3. Based on the spatial resolution of the rainfall, runoff, and underlying surface datasets described in Section 2.3.1, four modeling resolutions—1 km, 3 km, 5 km, and 10 km—were selected for distributed hydrological simulations. GDHF models were constructed at each resolution, followed by parameter calibration and validation at all ten hydrological stations. To assess the benefits of spatial refinement, an additional sub-watershed modeling scheme with fewer computational units was developed, consistent with the configuration used in the operational flood forecasting system of the Jialing River Basin (Zhang et al., 2022). An improvement index (IMP) was defined to quantify the performance enhancement of grid-based models relative to the sub-watershed scheme. The XGBoost model was then applied to evaluate the contribution of flood characteristics to the IMP metrics, identify key scale-sensitive controlling factors, and reveal the nonlinear effects of these factors on simulation

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

accuracy. Finally, flood events were classified based on these key factors, and the optional spatial resolutions under varying flood characteristics were statistically identified to support the selection of spatial resolution for large-scale hydrological modeling.

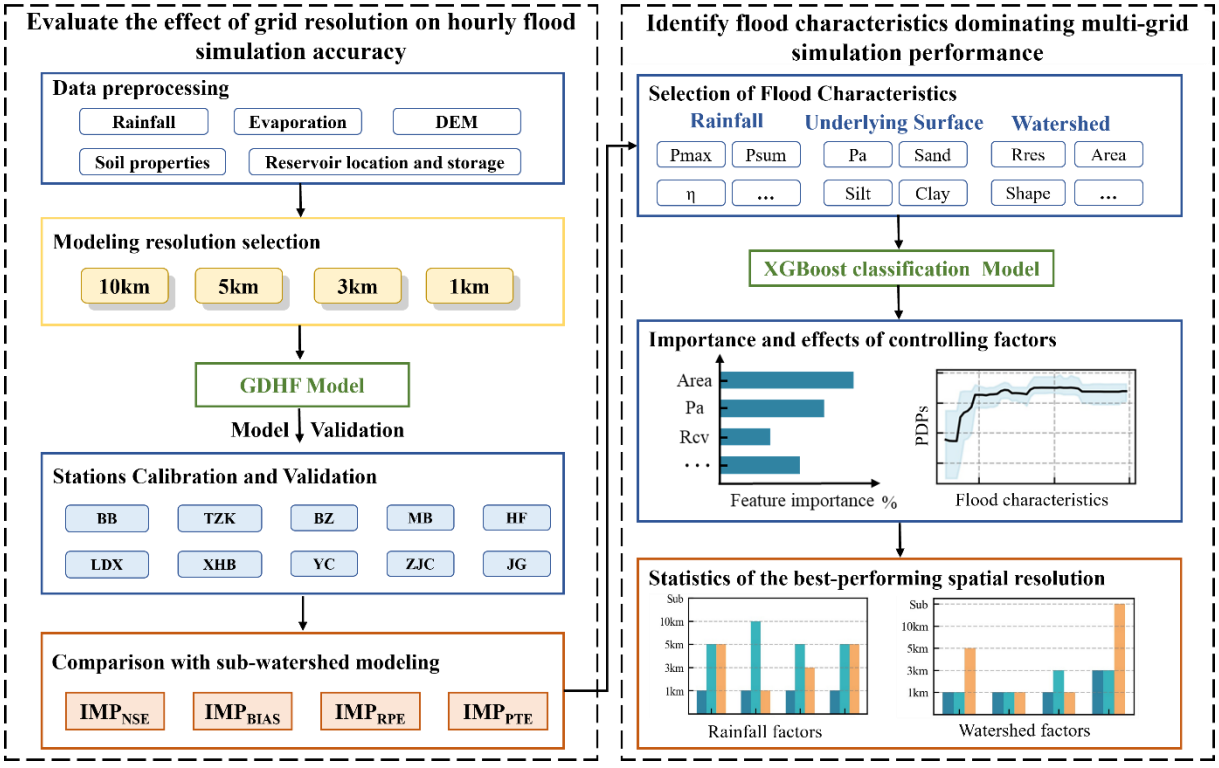


Figure 3. Experimental design on the impact of different spatial resolutions on the accuracy of hourly flood simulations.

2.3.1 Equations

The GDHF models at different spatial resolutions were calibrated using the NSGA-II algorithm. For each spatial resolution, an independent parameter set was calibrated separately for each hydrological station, using 60% of the flood events for calibration and the remaining 40% for validation. Model performance was evaluated using several statistical metrics, as summarized in Table 2. Relative error of runoff (BIAS) was used to assess the accuracy of the runoff generation process, while Relative error of flood peak (RPE) and Peak timing error (PTE) were employed to evaluate routing performance. Nash–Sutcliffe efficiency coefficient (NSE) was used to assess the overall model performance. In Table 2, Q_{sim} and Q_{obs} represent the simulated and observed discharge, respectively, and T_{sp} and T_{op} denote the simulated and observed peak flow times.

Table 2. Statistical metrics employed for the GDHF model calibration and assessment

Statistical metrics	Meaning	Equation	Perfect value

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

NSE	Nash-Sutcliffe efficiency coefficient	$NSE = 1 - \frac{\sum(Q_{sim} - Q_{obs})^2}{\sum(Q_{obs} - \overline{Q_{obs}})^2}$	1
BIAS	Relative error of runoff	$BIAS = \frac{\sum Q_{sim} - \sum Q_{obs}}{\sum Q_{obs}}$	0
RPE	Relative error of flood peak	$RPE = \frac{\max(Q_{sim}) - \max(Q_{obs})}{\max(Q_{obs})}$	0
PTE	Peak timing error	$PTE = T_{sp} - T_{op}$	0

To quantify the extent to which spatial-resolution refinement improves model performance, this study introduces the IMP, which measures the maximum gain achieved by grid-based models relative to sub-watershed modeling. IMP adopts a maximum-based statistic to characterize the upper bound of potential performance gains from spatial-resolution refinement, thereby preventing mean-based statistics from obscuring significant improvements at particular spatial resolutions (Tudaji et al., 2025). IMP_{NSE} represents the largest improvement in NSE by taking the difference between the highest NSE from all grid-based simulations and that of the sub-watershed model. IMP_{BIAS}, IMP_{RPE}, and IMP_{PTE} represent the greatest reductions in BIAS, RPE, and PTE, respectively, by comparing the minimum absolute BIAS, RPE, and PTE values from the grid-based simulations with those of the sub-watershed model:

$$IMP_{NSE} = Max\{NSE_{Grid,i}\} - NSE_{sub-watershed} \quad (3)$$

$$IMP_{BIAS} = |BIAS_{sub-watershed}| - Min\{|BIAS_{Grid,i}|\} \quad (4)$$

$$IMP_{RPE} = |RPE_{sub-watershed}| - Min\{|RPE_{Grid,i}|\} \quad (5)$$

$$IMP_{PTE} = |PTE_{sub-watershed}| - Min\{|PTE_{Grid,i}|\} \quad (6)$$

Where Grid and Sub-watershed refer to the grid-based and sub-watershed models.

2.3.2 Selection of Flood Characteristic Indicators

To identify flood characteristics most sensitive to model spatial resolution, fourteen indicators closely related to hydrological processes were selected from three categories: rainfall characteristics, underlying surface conditions, and watershed features. Descriptions are provided in Table 3, with detailed calculation methods in Appendix A.

Rainfall indicators include maximum rainfall intensity (P_{max}), total rainfall (P_{sum}), average rainfall intensity (P_{avg}), rainfall spatial non-uniformity coefficient (η), and temporal variation coefficient (R_{cv}). These capture the magnitude, intensity, and spatiotemporal variability of rainfall during each flood event. Underlying surface indicators include antecedent precipitation (P_a), reservoir retention capacity (R_{res}), and soil texture composition (*Sand*, *Silt*, and *Clay*). P_a reflects antecedent soil moisture before each flood event, R_{res} characterizes flood-regulation capacity, and soil texture variables represent infiltration and storage properties. Watershed characteristic indicators include watershed area (*Area*), watershed shape (*Shape*), average slope (*Slope*), and mean elevation (*Elev*), which describe the spatial scale, geomorphology, and topographic relief. Collectively, these indicators represent the dominant physical factors affecting runoff generation and routing.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Table 3. Meaning of 14 flood-characteristic indicators

Category	Criterion	Abbreviation	Unit	Explanation
Rainfall	Maximum rainfall intensity	P_{max}	mm/h	Maximum watershed-averaged rainfall intensity during the flood event
	Total rainfall	P_{sum}	mm	Total watershed-averaged rainfall amount for the flood event
	Average rainfall intensity	P_{avg}	mm/h	Ratio of total rainfall amount to total rainfall duration
	Rainfall spatial non-uniformity coefficient	η	-	Measures the spatial concentration or dispersion of rainfall
	Rainfall temporal variation coefficient	R_{cv}	-	Measures the variability of the rainfall time series
	Antecedent precipitation	P_a	mm	Reflects the antecedent soil moisture condition
Underlying Surface	Sand	Sand	%	Average sand fraction of the grid
	Silt	Silt	%	Average Silt fraction of the grid
	Clay	Clay	%	Average Clay fraction of the grid
	Reservoir retention capacity	R_{res}	mm	The reservoir's regulation capacity for flood within the watershed
Watershed	Watershed area	*Area*	km^2	The total area encompassed by the watershed
	Watershed shape factor	*Shape*	-	Reflects the shape characteristics of the watershed
	Watershed average slope	*Slope*	degree	Indicates the variation in terrain relief
	Watershed average elevation	*Elev*	m	Reflects the overall elevation of the watershed

2.3.3 XGBoost classification model

To examine how flood characteristics influence simulation results across different spatial resolutions, an XGBoost-based classification model was developed to identify key controlling factors most sensitive to spatial scale. XGBoost (Chen & Guestrin, 2016), a gradient-boosting decision tree algorithm, is widely used for hydrological applications due to its

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

efficiency and ability to capture nonlinear interactions (Niazkar et al., 2024; Fu et al., 2025). For binary classification, the model predicts the class probability by summing the outputs of multiple regression trees:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F \quad (7)$$

where $\hat{y}_i$ is the raw score for sample i, f_k is the k-th regression tree, and F is the space of all trees.

The XGBoost training procedure consisted of the following steps:

(a) Data preprocessing. 4 IMP metrics and 14 flood characteristics were used as model input variables. IMP values are affected by multiple spatial resolutions and are difficult to predict reliably. Therefore, the IMP metrics were converted into binary labels: label = 1 if spatial refinement yields a significant improvement (IMP_{NSE} > 0.10, IMP_{BIAS} > 5%, IMP_{RPE} > 5%, or IMP_{PTE} > 1 h), otherwise, label = 0 (Ekmekcioğlu and Koc, 2022).

(b) Model training. A site-specific leave-one-station-out validation strategy is employed to assess the model's spatial generalization ability. In each iteration, one hydrological station is used as an independent test set, while the remaining stations are used for model training and parameter optimization. This process is repeated for all stations, yielding 10 independent training–testing experiments. Within each training set, five-fold cross-validation is used for XGBoost hyperparameter tuning. This design ensures spatial independence between the training and test data, thereby effectively preventing spatial information leakage.

(c) Feature-effect interpretation. Partial dependence plots (PDPs) were used to visualize how each flood characteristic affects the IMP. For each plot, one feature was varied while all others were averaged over their empirical distributions. PDPs were averaged across 100 runs, and 95% confidence intervals were provided to indicate uncertainty.

2.3.4 Determination of the optimal spatial resolution for hydrological modeling

To systematically identify appropriate spatial scales under varying flood characteristics, this study employs three resolution-selection schemes based on the model performance. S1 selects the spatial resolution with the highest simulation accuracy within each characteristic interval; S2 identifies resolutions whose performance is within a predefined threshold of the optimal value; S3 determines appropriate resolutions using predefined accuracy criteria. Table 4 summarizes these three spatial-resolution selection schemes.

Table 4. Summary of the spatial resolution selection schemes employed in this study

Scheme	Selection criterion	Decision basis	Description
S1	Select the resolution with the highest simulation accuracy	Best value among NSE, BIAS, RPE, or PTE	Chooses the highest-accuracy resolution within each characteristic range

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

S2	Performance difference from the optimal value is below a threshold	Accuracy loss within the threshold (≤5% relative to the optimal value)	If multiple resolutions qualify, the coarser one is preferred
S3	Meets the predefined accuracy threshold	NSE > 0.80, or BIAS < 10%, or RPE < 10%, or PTE < 3 h	If multiple resolutions qualify, the coarser one is preferred; if none qualify, the best-performing resolution is selected

3 Hydrological modeling and evaluation at different spatial resolutions

The GDHF modeling at different spatial resolutions requires watershed unit delineation, spatial interpolation of rainfall, and discretization of underlying surface characteristics. The watershed unit delineation results are shown in Fig. A1, where the numbers of units for the sub-watershed scheme and for the 10 km, 5 km, 3 km, and 1 km spatial resolutions are 107, 1,662, 6,353, 17,283, and 152,243, respectively. Hourly rainfall observations were interpolated into cells of different resolutions using the Inverse Distance Weighting (IDW) method and used as rainfall inputs for the GDHF model.

The modeling process emphasizes spatial discretization of the underlying surface, including soil type, reservoir distribution, and river routing topology. Runoff-related parameters—such as field capacity, wilting point, and reservoir storage capacity—were derived from soil type, soil thickness, and the spatial distribution and storage capacities of small- and medium-sized reservoirs. Routing parameters (e.g., channel length, flow direction, and slope) were calculated from DEM data. Fig. 4 shows the spatial parameterization of the GDHF model using the 10 km resolution as an example.

As illustrated in Fig. 4(a), Loam and Silty Loam are the dominant soil types in the watershed. Fig. 4(b) presents the soil moisture capacity ($S_0 + U_0$) of the surface and lower layers, estimated from soil type and soil thickness, with values ranging from 130 to 170 mm and showing pronounced spatial variability. Fig. 4(c) shows the flow directions and river network topology extracted from the DEM, indicating that 1,662 grid cells ultimately drain to the watershed outlet. Fig. 4(d) illustrates the influence of reservoir regulation, which is weak in upstream areas but becomes significant in downstream regions.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

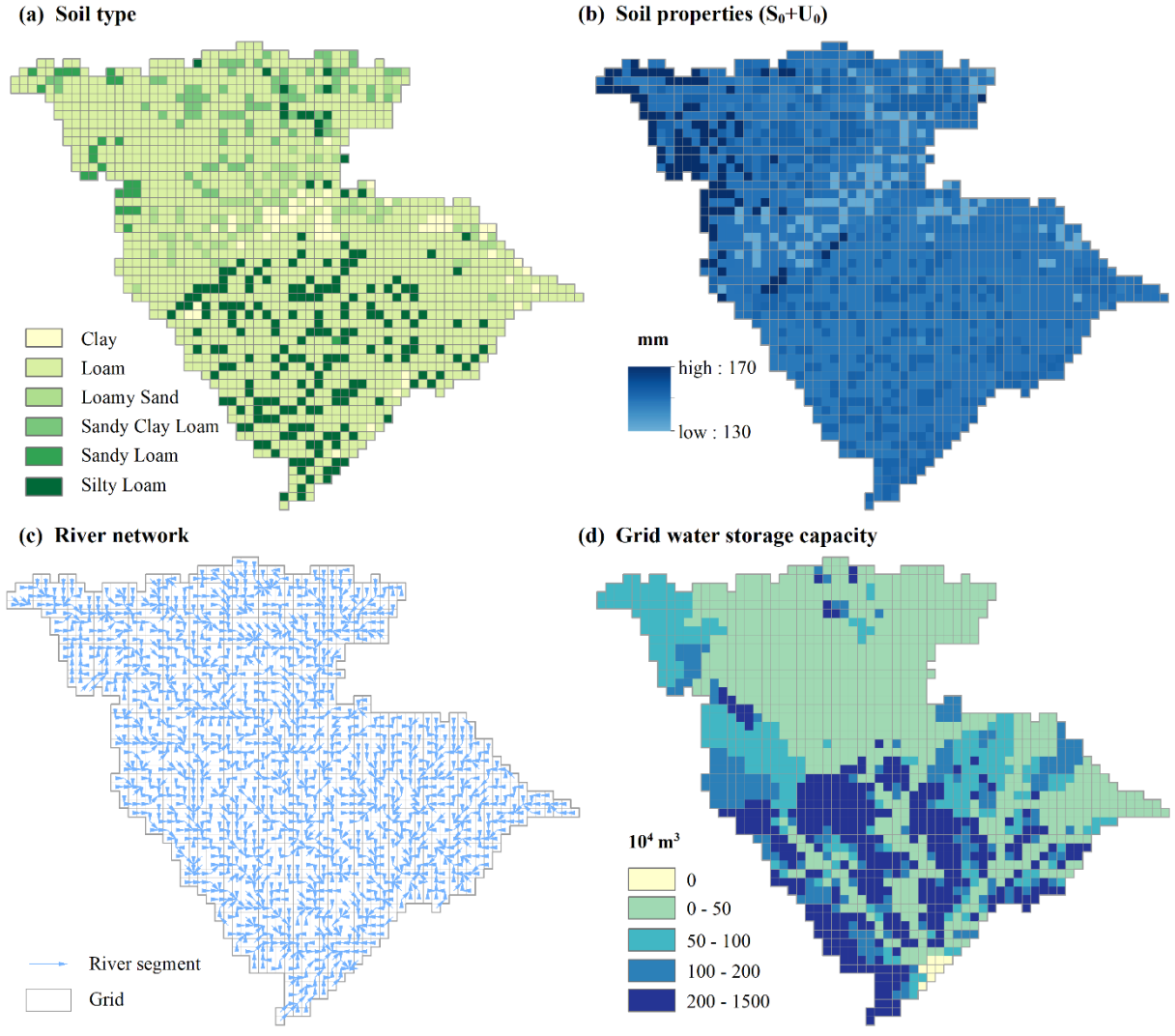


Figure 4. Spatial distribution of (a) the soil types, (b) soil properties, (c) river network, and (d) grid water storage capacity.

4 Results and Discussion

4.1 Comparison of flood simulation accuracy at different spatial resolutions

230 Based on the evaluation metrics described in Section 3, the performance of flood simulations under different spatial modeling resolutions was compared for both the calibration and validation periods across the ten stations (Fig. 5). The evaluation metrics |BIAS|, |RPE|, and |PTE| represent the average absolute errors of all flood events at each station.

Fig. 5 shows that flood simulation accuracy varies substantially across spatial resolutions. Among all configurations, the 1 km grid performs the best: NSE values exceed 0.80 at all stations, while |BIAS| and |RPE| remain around 10% and 15%,

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

respectively, and |PTE| consistently stays below 5 h. This indicates that the 1 km resolution effectively captures the runoff-generation and flow-routing processes of flood events. In contrast, the sub-watershed model, which contains the fewest computational units, exhibits the lowest performance, with NSE values dropping to approximately 0.60 at several stations and flood-volume and peak-flow errors exceeding 20%.

The sensitivity of simulation performance to spatial resolution differs markedly across station types. For stations with large control areas (such as BB), differences in simulation accuracy among the 1 km, 3 km, 5 km, and even 10 km grids are minimal, and the sub-watershed model can still reproduce the major flood characteristics reasonably well. In contrast, for stations with smaller control areas (BZ, MB, HF, YC, ZJC, and JG), refining the spatial resolution substantially enhances the representation of runoff generation and flow routing, thereby significantly improving simulation accuracy.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

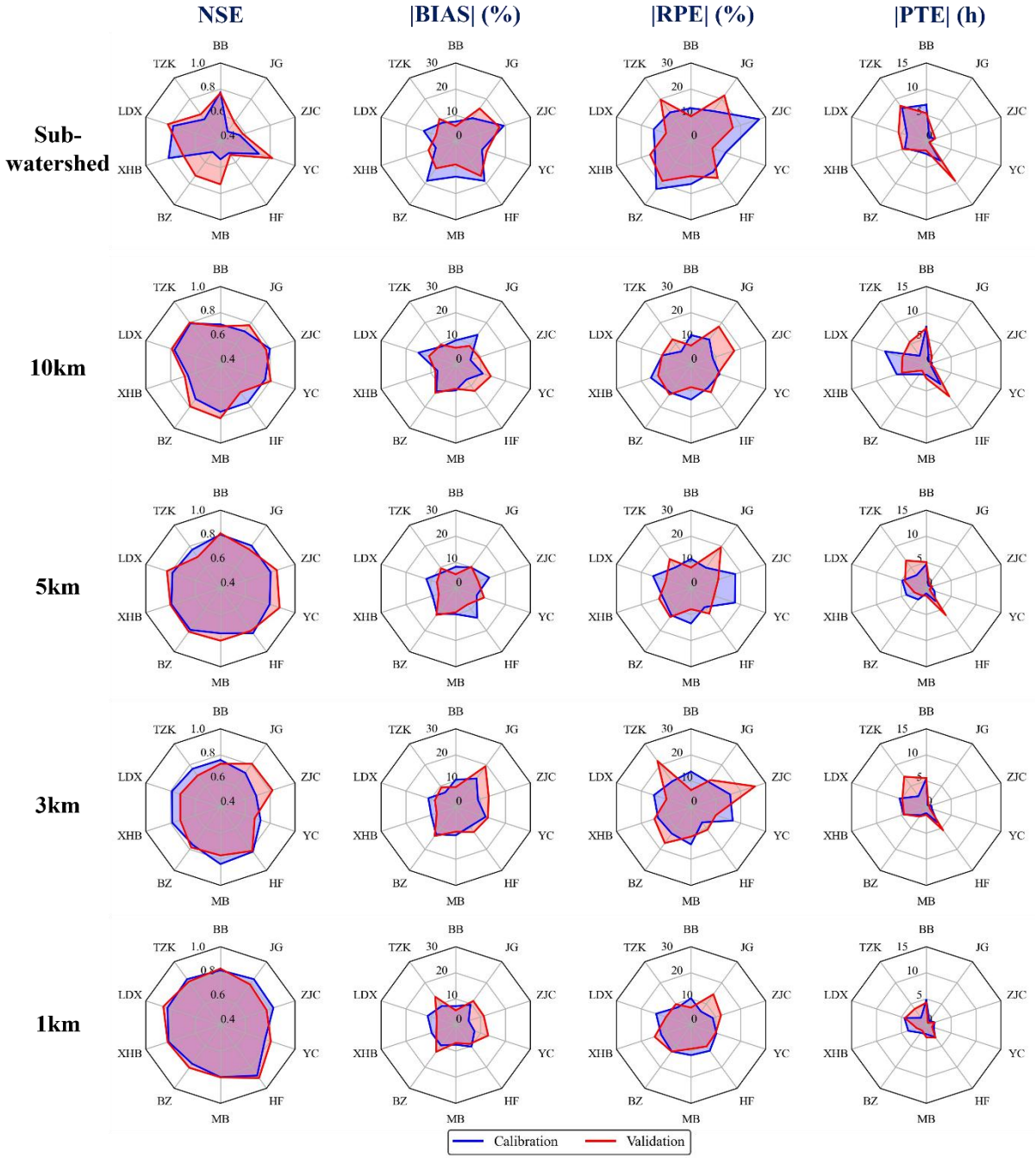


Figure 5. Evaluation metrics for simulating hourly flood events across different spatial scales during calibration and validation periods.

To further assess spatial resolution effects for stations of different areas, the stations were classified into two categories: control stations and small-to-medium watersheds. The control stations (BB, TZK, LDX, and XHB) are located at watershed

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

outlets or major tributaries, with control areas of 28,826–157,041 km². The remaining stations represent small-to-medium watersheds with control areas of 233–2,707 km².

Fig. 6(a) shows that for small-to-medium watersheds, NSE values increase substantially as the spatial resolution is refined. Compared with coarse-resolution modeling (sub-watershed, 10 km, and 5 km), the 3 km and 1 km grids capture the rising and recession limbs of the hydrograph much more accurately. In contrast, for control stations, refining the spatial resolution results in only marginal improvements in NSE. Figs. 6(b), 6(c), and 6(d) further indicate that for small-to-medium watersheds, finer spatial resolution leads to substantial improvements in |BIAS|, |RPE|, and |PTE|, whereas for control stations, refining the spatial resolution yields almost no improvement in simulated flood volume, peak flow, or peak timing.

This contrast reflects fundamental differences in hydrological nonlinearity across watershed areas. Small watersheds are highly sensitive to localized rainfall, steep gradients, and rapid flow routing, resulting in strong nonlinear responses. In large watersheds, however, floods originate from multiple sub-regions, and nonlinear processes become smoothed during flow aggregation, reducing the overall nonlinearity of the flood response. Consistent with these findings, Jiang et al. (2025) also reported that higher-resolution CaMa-Flood models improve daily peak-flow simulation, with the most pronounced improvements occurring in smaller tributaries.

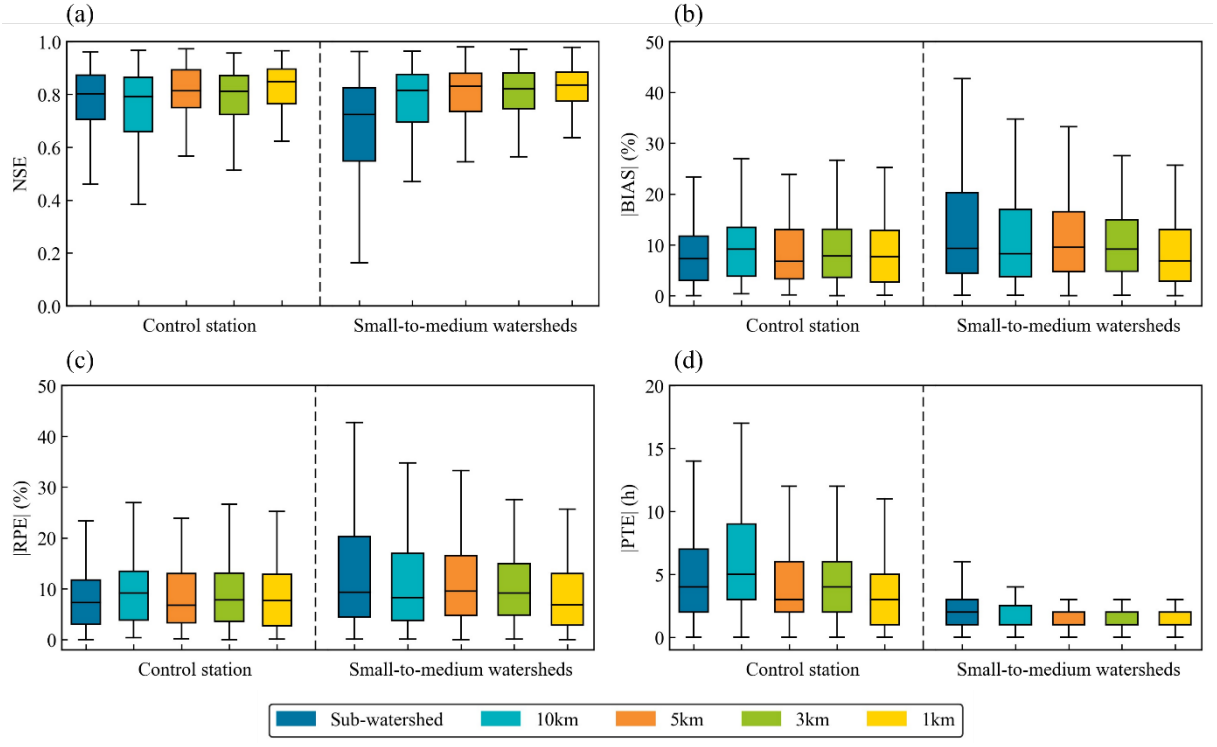


Figure 6. Comparison of flood simulation results across various spatial resolutions for stations with different areas.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

4.2 Comparison of computational time at different spatial scales

To compare the computational time under different spatial resolutions, the flood event from August 22 to 29, 2020, was selected. After applying a 10-day warm-up period, the effective simulation spanned 17 days, consisting of 408-time steps. All computations were performed on a laptop equipped with an 11th Gen Intel® Core™ i7-11800H processor (8 cores and 16 threads). The corresponding computational times are summarized in Table 5.

The results show that as spatial resolution is refined, the number of grid cells increases exponentially, leading to a sharp rise in computational time. The sub-watershed and 10 km simulations require only about 13 s and 1 min 11 s, respectively, whereas the 1 km simulation requires approximately 18 min 56 s. Moreover, real-time rolling forecasting requires preparing large volumes of hydrological input data—such as rainfall interpolation and initial soil moisture conditions—which further increases computational burden. Therefore, applying a 1 km grid resolution to a large watershed significantly increases computational demand, making it difficult to meet the efficiency requirements of real-time flood forecasting.

Table 5. Computational time for simulating a single flood event at different spatial resolutions

Spatial resolution	Sub-watershed	10km	5km	3km	1km
Number of cells	107	1662	6353	17283	152243
Calculation Time	13s	1 min 11 s	1 min 40 s	2 min 47 s	18 min 56 s

4.3 Identifying controlling factors of spatial resolution with XGBoost

Flood characteristics identified in Section 2.3.2 as influential to runoff generation and routing were examined as potential controls for determining the appropriate spatial scale. To ensure their robustness, we analyzed their relationships with the IMP metrics. Figure 7 presents the Spearman correlation coefficients (ρ) and significance levels for 14 flood characteristics and 4 IMP metrics. The sign of ρ indicates positive or negative correlation, values closer to ±1 represent stronger monotonic relationships, and values near 0 indicate weak associations. Asterisks denote the corresponding significance levels.

The results show that P_{max} and P_{avg} are strongly positively correlated (ρ = 0.90), indicating close consistency between maximum and average rainfall intensity. Thus, either can serve as the representative rainfall variable in the subsequent XGBoost model. The variable η also exhibits significant positive correlations with P_{max} and P_{avg} (ρ = 0.78 and 0.80). Regarding underlying surface characteristics, *Area* shows strong negative correlations with P_{max}, P_{avg}, and η (ρ = –0.82, –0.83, and –0.79), suggesting that as *Area* increases, rainfall intensity decreases and rainfall spatial variability increases. *Elev* is negatively correlated with *Silt* (ρ = –0.69) and positively correlated with *Slope* (ρ = 0.59). R_{res} also shows a significant negative correlation with *Slope* (ρ = –0.73), indicating that reservoir regulation weakens in higher-elevation watersheds.

Although the ρ values between these flood characteristics and four IMP metrics are not particularly high, many characteristics—such as *Area* and several rainfall- and surface-related indicators—still exhibit statistically significant relationships with IMP. The strong associations between these characteristics and especially IMP_{NSE} support the validity of

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

the selected potential driving factors and provide a solid foundation for identifying key features in subsequent analyses.

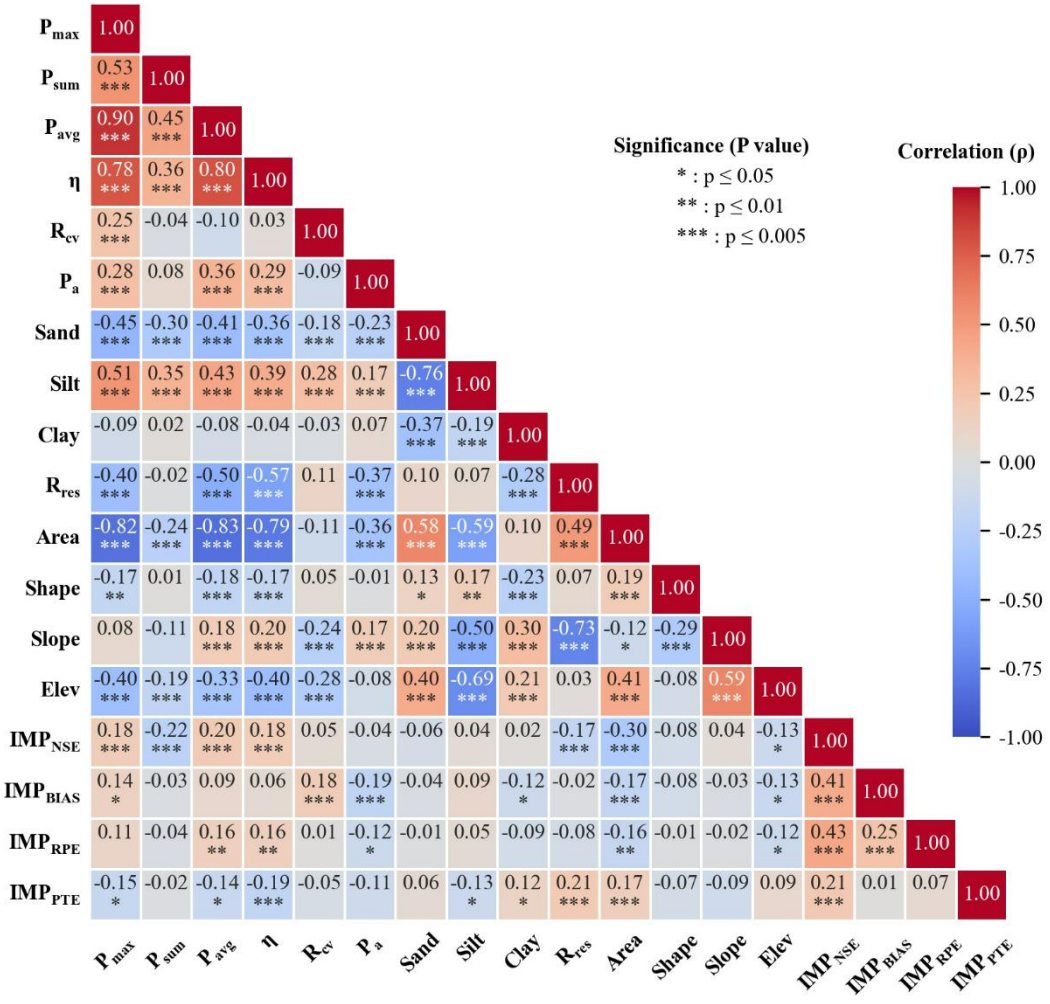


Figure 7. Spearman correlation analysis results between 14 flood characteristics and 4 IMP metrics.

To examine how potential driving factors influence flood simulations across different spatial scales, this study employs an XGBoost model to identify key scale-sensitive factors, with model configurations detailed in Section 2.3.3. Because the ρ between P_{max} and P_{avg} exceeds 0.90, indicating strong multicollinearity, P_{avg} was excluded, and P_{max} was retained as the representative rainfall intensity indicator. The trained XGBoost model achieved average recall values of 78.3%–97.8% on the training sets and 64.9%–73.0% on the test sets across all IMP metrics, demonstrating good fitting performance and reasonable generalization. This confirms its suitability for identifying the dominant factors influencing flood simulation outcomes.

Fig. 8 presents the mean contribution rates of the driving factors and their standard deviations across the IMP metrics. For

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

the IMP_{NSE} metric (Fig. 8a), *Area*, P_{sum}, and *Slope* are the primary contributors, explaining 40.9% of the model output, with *Area* showing the highest importance and thus serving as the key controlling factor. For the IMP_{BIAS} metric (Fig. 8b), P_a, *Clay*, and R_{cv} contribute most significantly (35.4%), indicating that antecedent soil moisture, soil clay content, and maximum rainfall intensity—three variables strongly linked to runoff generation—jointly govern the IMP_{BIAS} response. For the IMP_{RPE} metric (Fig. 8c), P_{max}, P_{sum}, and *Elev* are the dominant features, accounting for 32.6% of the contribution. For the IMP_{PTE} metric (Fig. 8d), *Slope*, R_{res}, and *Clay* emerge as the main drivers, with a combined contribution of 26.3%. These results highlight that rainfall characteristics, together with *Elev*, *Slope*, R_{res}, and *Clay*, are critical factors affecting peak discharge and time to peak. Overall, the XGBoost analysis reveals that the flood characteristics controlling improvements in multi-grid simulation performance differ substantially across the various IMP evaluation metrics.

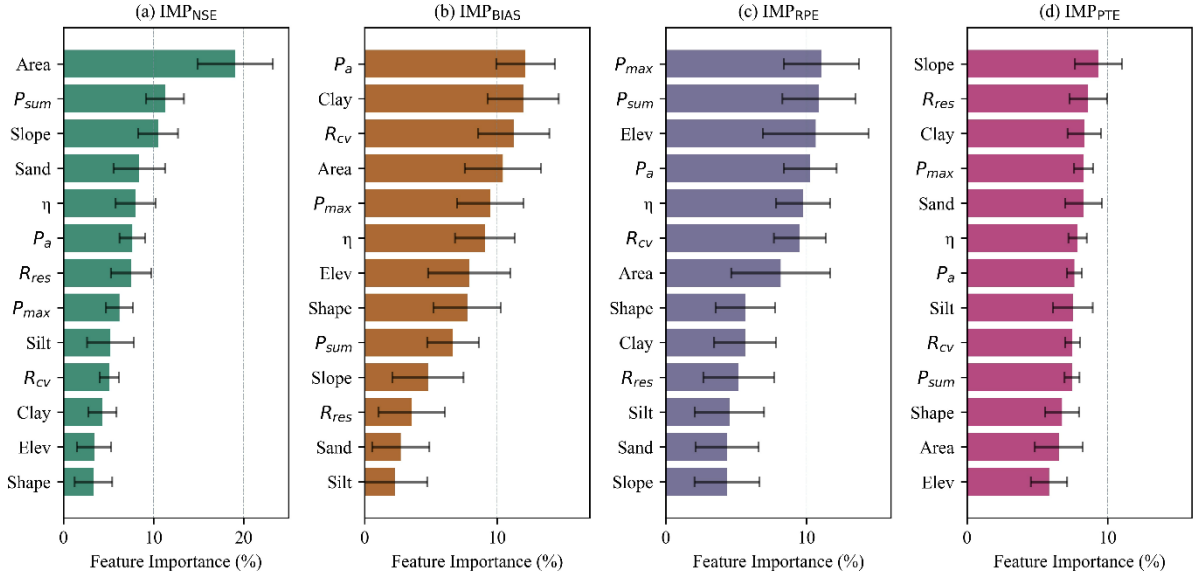


Figure 8. Importance ranking and contribution rates of flood characteristics based on the XGBoost model.

4.4 Nonlinear effects of controlling factors on flood simulation accuracy

To further clarify how key flood characteristics influence the IMP values, this study uses PDPs to illustrate the relationships between each characteristic and the IMP, as shown in Fig. 9.

For IMP_{NSE}, Fig. 9(a) shows that smaller watersheds are more likely to benefit from spatial refinement, and the probability of improvement declines progressively as watershed area increases. In addition, Figs. 9(b) and 9(c) indicate that lower P_{sum} and lower *Slope* values are associated with higher probabilities of IMP_{NSE} improvement. For IMP_{BIAS}, Fig. 9(d) identifies P_a as a key determinant: smaller P_a values increase the likelihood of improving IMP_{BIAS}. Fig. 9(e) further shows that lower *Clay* content also favors performance gains. Moreover, Fig. 9(f) demonstrates that larger R_{cv}—reflecting more pronounced temporal variability of rainfall—corresponds to an increased probability of improving IMP_{BIAS}.

For IMP_{RPE}, Fig. 9(g) shows that higher P_{max} enhances the likelihood of improvement, while Figs. 9(h) and 9(i) indicate that

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

lower P_{sum} and lower *Elev* likewise increase the probability of IMP_{RPE} improvement. For IMP_{PTE}, Figs. 9(j) and 9(k) reveal that watersheds with gentler *slopes* and higher R_{res} (i.e., stronger reservoir regulation effects) are more likely to achieve reductions in peak-timing error through spatial refinement. However, Fig. 9(l) shows that the influence of *Clay* on IMP_{PTE} fluctuates without a clear monotonic trend.

Overall, *Area* exerts the most pronounced nonlinear effect on the probability of improvement for IMP_{NSE}. Rainfall characteristics—particularly P_{sum}, P_{max}, and R_{cv}—play critical roles in determining improvement probabilities for IMP_{NSE}, IMP_{BIAS}, and IMP_{RPE}. In addition, several underlying surface characteristics, including P_a, *Slope*, *Clay*, *Elev*, and $R_{res,}$ also contribute to variations in the effectiveness of IMP improvements.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

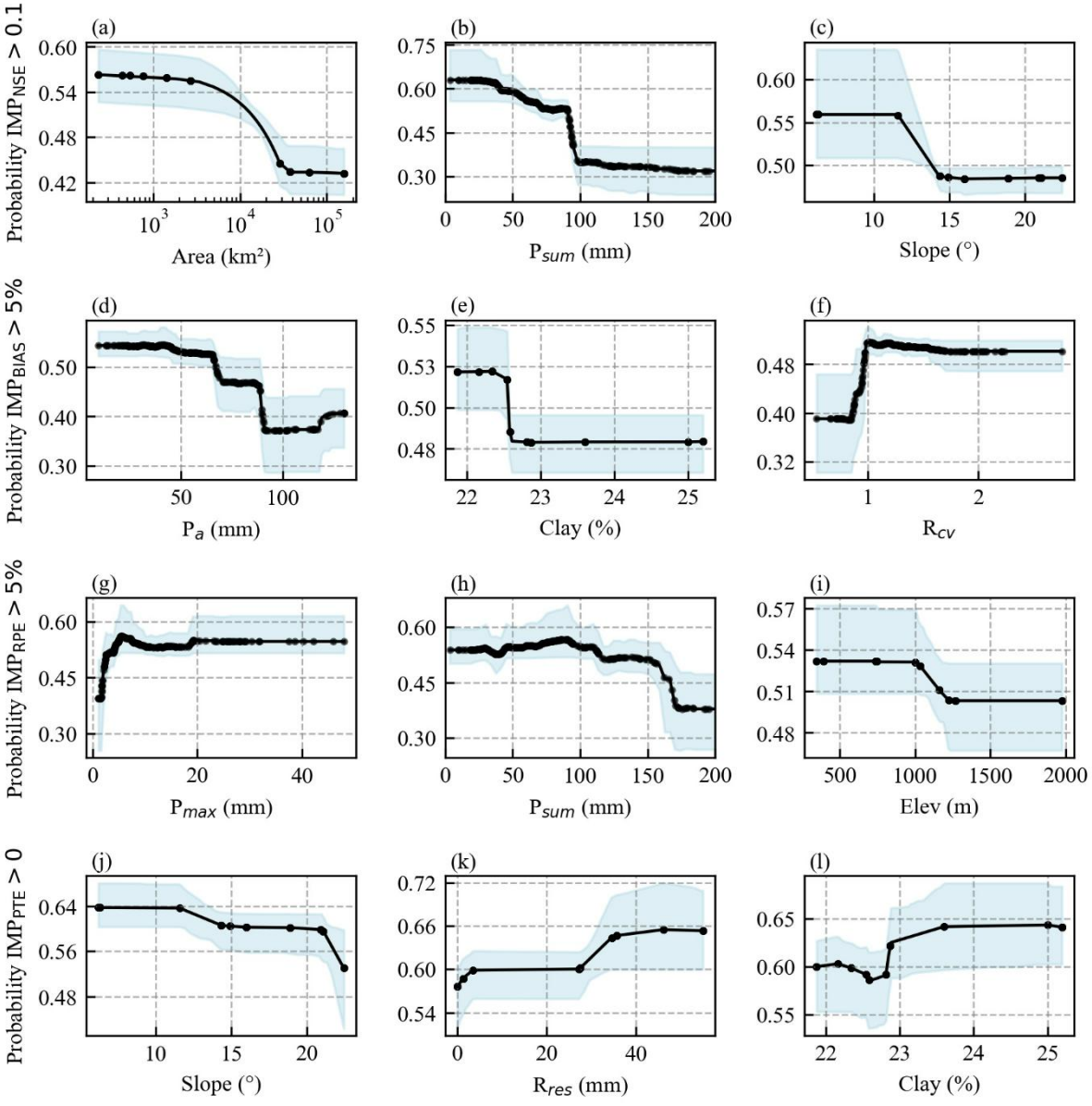


Figure 9. PDPs of the predicted probabilities of improvement in four IMP metrics based on different flood characteristics.

335 To validate the nonlinear influence of flood characteristics on the IMP revealed by the PDPs, flood events were classified according to their characteristic values, and the IMP metrics were statistically analyzed across categories, as shown in Fig. 10.

Fig. 10(a) shows that when the *Area* is small (<10^4 km²), particularly below 10^3 km², most flood events yield positive IMP_{NSE} values, indicating that small and medium-sized watersheds are most responsive to spatial refinement. As the 340 watershed area increases, the median IMP_{NSE} approaches zero, indicating that spatial refinement offers limited NSE

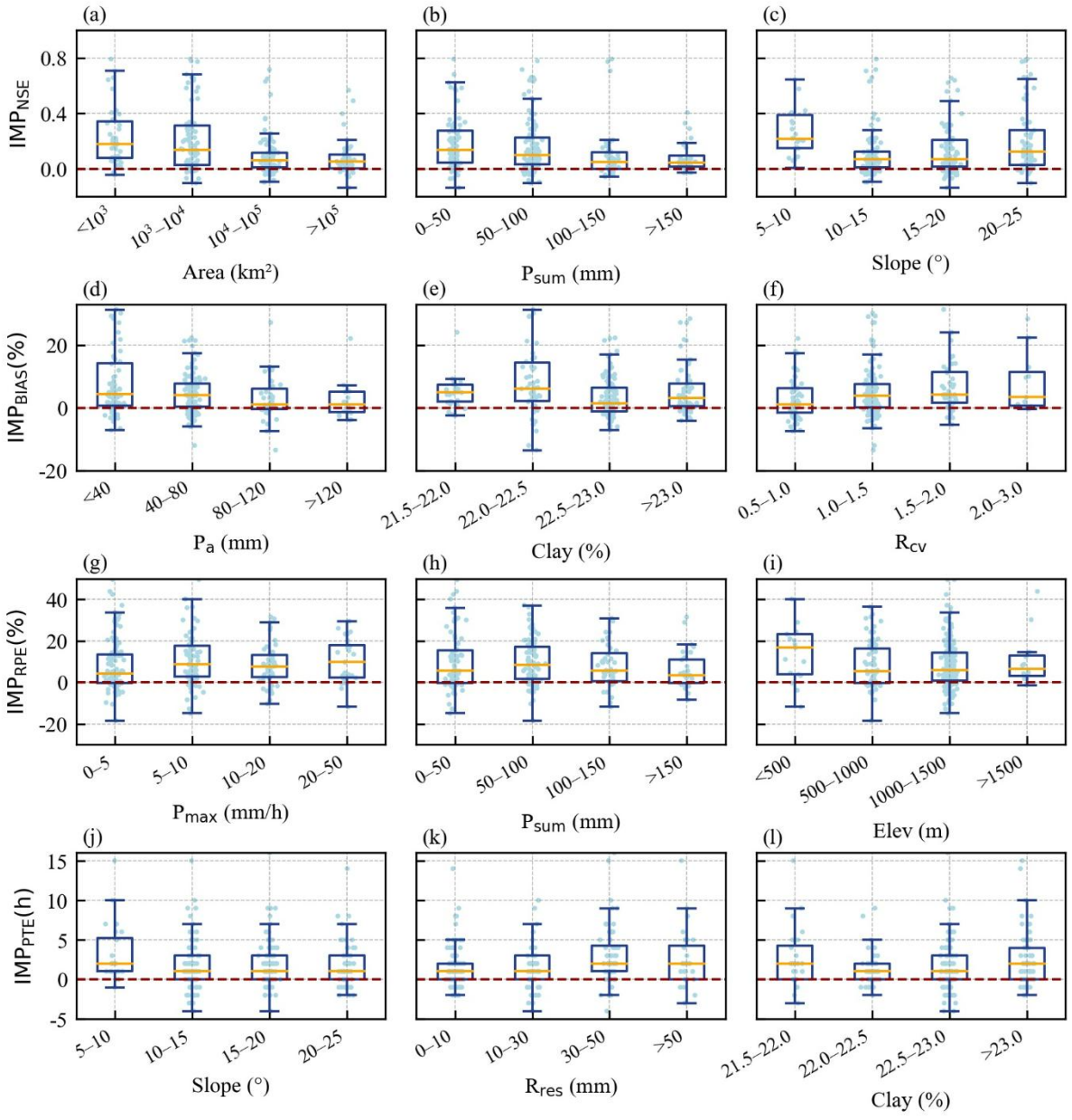
https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

improvement in large watersheds. Thus, *Area* is the dominant factor governing whether NSE can benefit from finer spatial resolution. Fig. 10(b) further shows that IMP_{NSE} improves most substantially when P_{sum} <100 mm, whereas flood events with larger rainfall amounts (P_{sum} >100 mm) exhibit much weaker sensitivity. Similarly, Fig. 10(c) indicates that when *Slope* <10°, IMP_{NSE} values are almost entirely positive, suggesting that watersheds with gentler terrain benefit more from spatial refinement.

For IMP_{BIAS}, Fig. 10(e) demonstrates a marked improvement in IMP_{BIAS} when P_a <80 mm—especially <40 mm—highlighting the nonlinear control of antecedent soil moisture on flood-volume simulations. Fig. 10(d) shows that when *Clay* <22.5%, most events have positive IMP_{BIAS} values, implying that low clay content favors reductions in flow-volume bias. As shown in Fig. 10(f), the improvement in IMP_{BIAS} is particularly significant when the rainfall temporal variability coefficient (R_{cv}) exceeds 1.5.

Figure 10(g) suggests that IMP_{RPE} varies little across P_{max} categories, whereas Fig. 10(h) shows that IMP_{RPE} improves noticeably when P_{sum} <100 mm. Figure 10(i) reveals that low-elevation basins tend to achieve greater RPE improvement, with median IMP_{RPE} decreasing as *Elev* increases. For IMP_{PTE}, Fig. 10(j) shows that *Slope* <10° yields the greatest reduction in peak-timing error. Fig. 10(k) further indicates that larger R_{res} values—representing stronger reservoir regulation—are associated with greater improvements in IMP_{PTE}, suggesting that regulated basins benefit more from spatial refinement than natural watersheds. Finally, Fig. 10(l) shows no clear trend between IMP_{PTE} and *Clay*, implying that peak-timing errors are influenced primarily by terrain and reservoir regulation rather than *Clay* content. Overall, flood characteristics impose distinct and nonlinear controls on the effectiveness of spatial resolution refinement.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026



360 **Figure 10. Boxplots of IMPNSE, IMPBIAS, IMPRPE, and IMPPTE metrics under different flood characteristics.**

4.5 Statistical analysis of optimal spatial resolution based on controlling factors

The analysis in Section 4.4 indicates that, under specific flood-characteristic conditions, refining the spatial resolution can substantially improve flood-simulation accuracy compared with sub-watershed modeling. Building on this, the present section further compares model performance across different spatial resolutions to identify the optimal spatial scale under

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

various controlling-factor scenarios. Fig. 11 illustrates how NSE, BIAS, RPE, and PTE vary with flood characteristics at different spatial resolutions.

Overall, within the characteristic ranges identified in Section 4.4, finer spatial resolutions generally yield higher simulation accuracy than sub-watershed modeling, as reflected by increased NSE and decreased BIAS, RPE, and PTE values. The improvements are particularly pronounced for small watersheds (*Area* $<10^4$ km^2), gentle terrain (*Slope* $< 10^{\circ}$), or low cumulative rainfall (P_{sum} < 100 mm), where the 1 km, 3 km, and 5 km grids all outperform the sub-watershed approach. In contrast, for large basins (*Area* > 10^4 km^2), wetter antecedent soil conditions (P_a > 80 mm), or other conditions characterized by weaker flood nonlinearity, differences among spatial resolutions diminish, and spatial refinement offers limited benefits.

Although the 1 km grid generally achieves the best performance across all four metrics, Figure 11 also shows that in many characteristic ranges, the performance differences between resolutions are relatively small. This suggests that higher spatial resolutions do not necessarily yield proportional improvements. Therefore, spatial refinement should be applied selectively, guided by the degree of flood-process nonlinearity and the sensitivity of simulations to local rainfall and underlying-surface heterogeneity.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

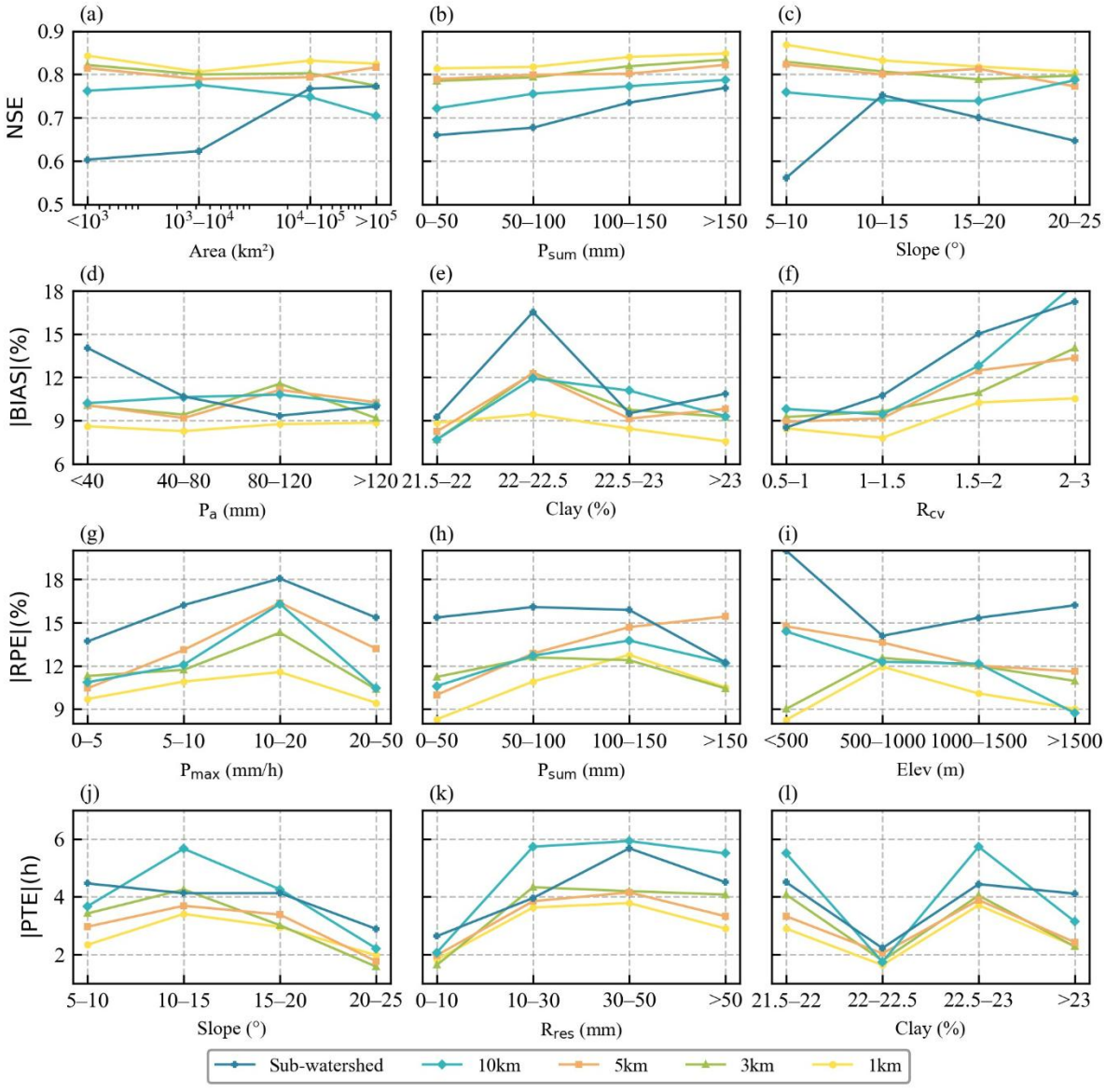


Figure 11. Variations of flood simulation accuracy with flood characteristics under different spatial resolutions.

To systematically identify appropriate spatial scales under varying flood characteristics, this study employs three resolution-selection schemes (Section 2.3.4) based on the performance patterns observed in Fig. 11. Fig. 12 presents the appropriate spatial resolutions for different flood characteristic categories under the three schemes.

The statistics from S1 show that fine grids (1 km and 3 km) generally perform best across most characteristic intervals. Nonetheless, notable exceptions exist: when the soil clay content falls within 21.5–22%, or when basin elevation exceeds 1500 m, the 10 km grid becomes the preferred resolution. This indicates that under these specific conditions, coarser grids

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

may in fact yield higher accuracy for certain performance metrics.

For the NSE metric, S2 tends to select 3–5 km grids across multiple characteristic ranges. This is because, under the criterion that a resolution is acceptable when its relative accuracy loss does not exceed 5%, medium-resolution grids can already achieve NSE performance very close to that of the optimal resolution. In contrast, for the BIAS, RPE, and PTE metrics, the spatial resolutions recommended by S2 are largely identical to those selected by S1. This indicates that, under these characteristic conditions, the performance gaps between the non-optimal resolutions and the optimal one remain relatively large, making it difficult for them to satisfy the 5% accuracy-loss threshold.

For S3, the NSE metric tends to favor the 5 km spatial resolution because this resolution already meets the predefined accuracy criterion (NSE > 0.80), making further refinement unnecessary. For BIAS, when antecedent soil moisture is high (P_a > 80 mm), clay content is relatively large, or rainfall exhibits a more uniform spatial distribution (R_{cv} < 1.5), S3 tends to recommend coarser resolutions (such as 10 km or Sub-watershed). These resolutions are sufficient to satisfy the accuracy threshold (BIAS < 10%), and therefore, higher spatial resolutions are not required. In contrast, for the RPE and PTE metrics, coarse grids often fail to meet the predefined accuracy requirements (RPE < 10%, PTE < 3 h). As a result, the S3 selections align closely with those of S1, more frequently recommending finer resolutions (1–3 km).

In summary, selecting spatial resolutions based on flood characteristics not only ensures reliable flood-forecasting accuracy but also avoids unnecessary computational costs associated with blindly refining model resolution, thereby providing a scientific basis for determining appropriate spatial scales in large-watershed flood forecasting.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

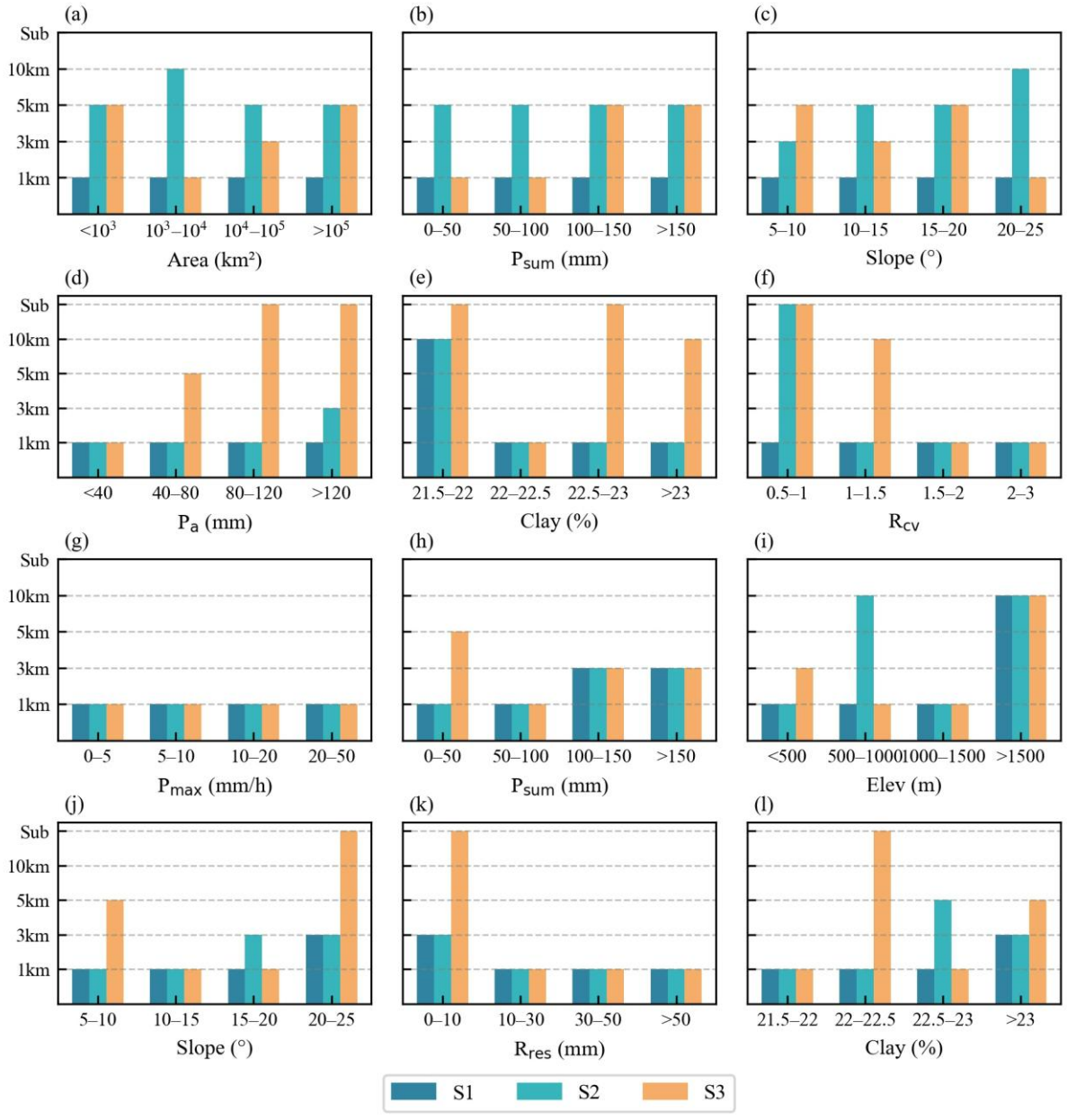


Figure 12. Spatial resolution recommendations for different flood characteristic categories based on the three schemes.

4.6 Effect of rainfall gauge density on flood simulation accuracy

To evaluate the effectiveness of refining model spatial resolution under sparse rainfall observations, this section reduces the number of rain gauges by 25%, 50%, and 75% and examines the influence of gauge density on flood-simulation accuracy across multiple spatial resolutions. To avoid representation issues associated with random gauge removal, the Thiessen

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

polygon method was used to compute the weight of each station. Rain gauges were then ranked in descending order of weight, and those with lower weights were removed first to ensure that the remaining stations best represented the spatial distribution of rainfall across the study area.

Given that the control stations are less sensitive to changes in spatial modeling resolution than small and medium-sized watersheds (Fig. 6), this section focuses on all small and medium-sized watersheds for analysis. Fig. 13 shows the flood simulation accuracy across different spatial resolutions under varying rainfall station removal ratios (0%, 25%, 50%, and 75%). The results show that when rainfall-gauge density is high (e.g., no removal or only 25% removal of gauges), flood simulation accuracy increases significantly with finer spatial resolution, and the median NSE is higher for fine grids (1–3 km). However, when rainfall-gauge density is substantially reduced (with 50% or 75% of the gauges removed), the differences in simulation accuracy among different spatial scales become much smaller, and spatial refinement provides only limited improvement to flood-simulation accuracy. This indicates that under sparse gauge conditions, even fine-resolution grids cannot substantially improve simulation results. Therefore, when selecting the spatial resolution for modeling, the spatial representativeness of the rainfall data must be carefully considered.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

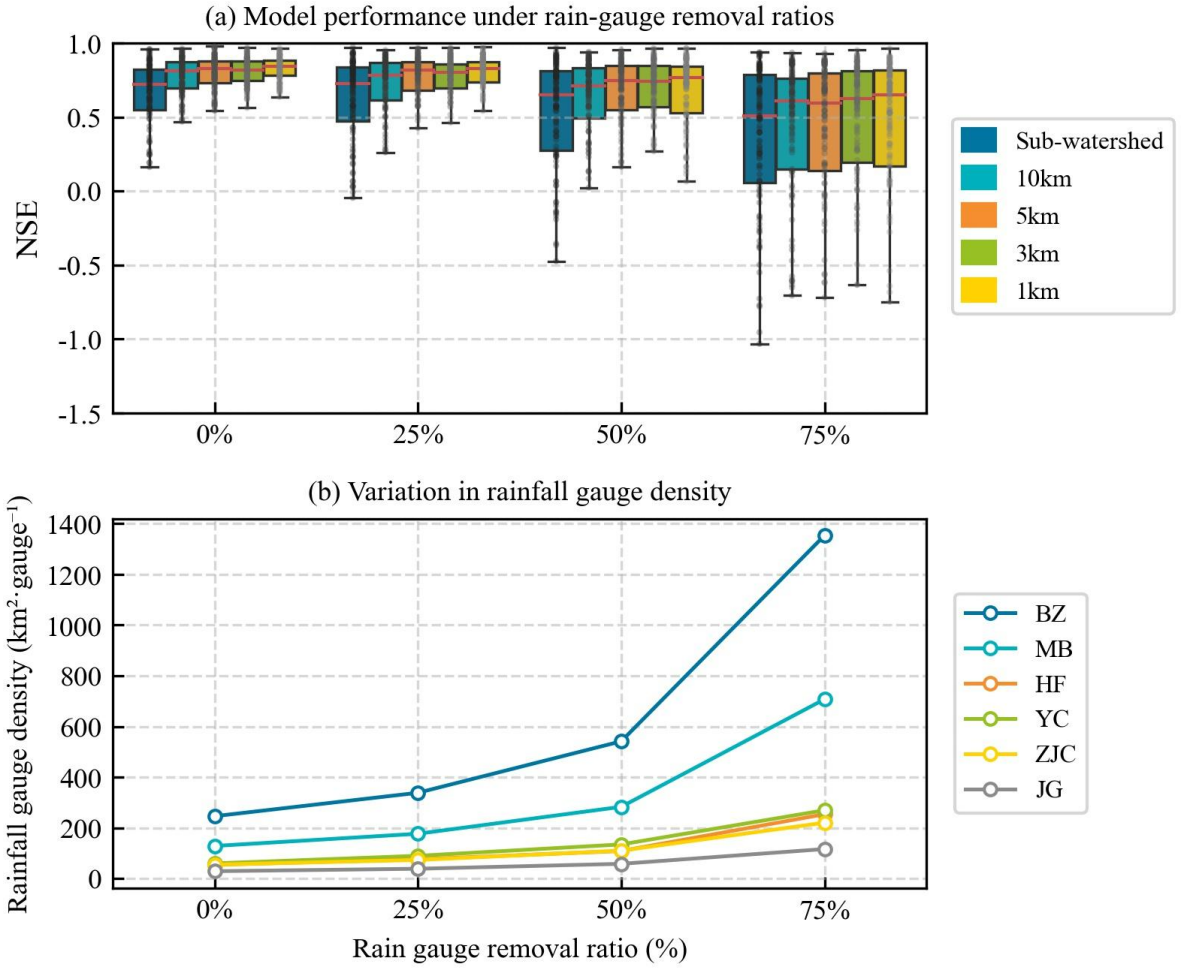


Figure 13. Flood simulation accuracy across different spatial scales under varying rainfall station removal ratios (0%, 25%, 50%, and 75%).

425 Building on this analysis, three representative flood events with notable accuracy gains from spatial-resolution refinement were selected to further assess whether fine-resolution grids (1 km) remain effective under reduced rainfall-gauge density. The resulting changes in rainfall spatial distribution and flood-simulation performance across different gauge densities are shown in Fig. 14.

Fig. 14(a) shows that as the gauge removal ratio increases from 0% to 75% (density reduced to 707.5 $km^2 \cdot gauge^{-1}$), the
430 rainfall field becomes increasingly uniform, its overall magnitude weakens, and the original storm center disappears. When the gauge network is relatively dense (0%, 25%, and 50% removal), the interpolated rainfall retains reasonable spatial structure, allowing the 1 km grid to maintain high simulation accuracy (NSE_{1km} = 0.86, 0.79, and 0.73). However, with 75%-gauge removal, the representativeness of the rainfall field deteriorates sharply, leading to a substantial decline in accuracy (NSE_{1km} = −0.64).

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Figs. 14(b) and (c) exhibit the same pattern: decreasing gauge density causes the rainfall field to deviate more from the true distribution, resulting in reduced simulation accuracy across resolutions. These results highlight that the benefits of spatial-resolution refinement are highly dependent on the spatial representativeness of rainfall input. Under sparse or poorly representative gauge networks, finer grids offer limited improvement and cannot compensate for deficiencies in rainfall observations. Therefore, when determining an appropriate modeling resolution, the spatial coverage and reliability of rainfall data should be evaluated as a priority rather than relying solely on spatial refinement to enhance model performance.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

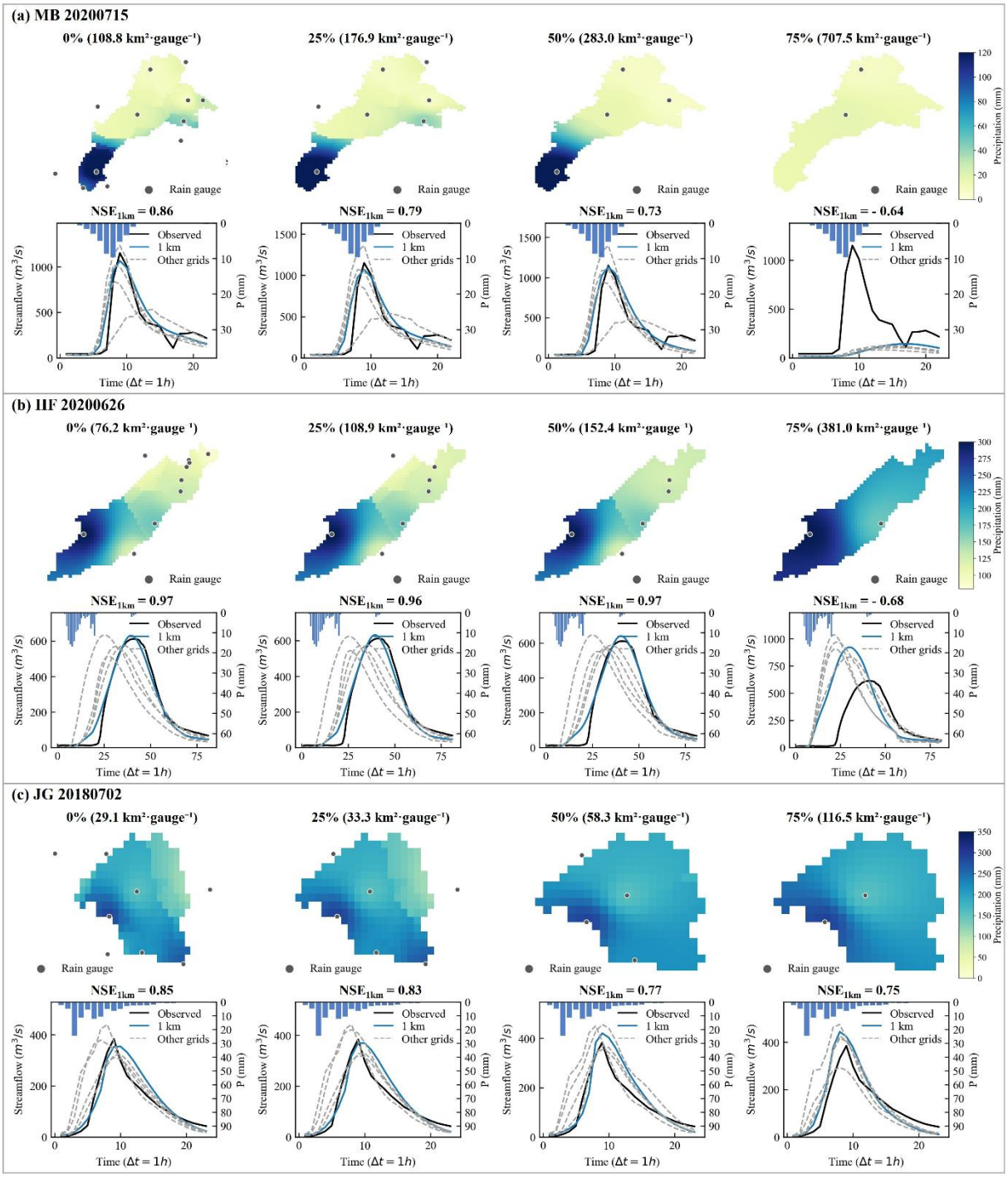


Figure 14. Effects of rainfall-gauge density on rainfall spatial distribution and flood simulation performance under different spatial resolutions.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

5 Conclusions

To investigate the influence of spatial scale on flood-simulation accuracy in large river basins, this study developed the GDHF model for the Jialing River basin at five resolutions: 1 km, 3 km, 5 km, 10 km, and sub-watershed. An XGBoost model was then used to identify key flood characteristics that govern the choice of spatial resolution. Based on these results, the study determined the optimal spatial resolution for different flood-characteristic ranges and further examined how rainfall-gauge density constrains the benefits of spatial refinement. The main conclusions are as follows:

(1) Multi-resolution simulations show that the 1 km grid consistently performs best, with NSE > 0.80 at all stations, |BIAS| and |RPE| around 10% and 15%, respectively, and |PTE| < 5 h. However, for stations controlling large drainage areas, additional spatial refinement yields only marginal accuracy gains. In contrast, for small- and medium-sized watersheds, finer grids markedly improve the representation of runoff generation and routing processes.

(2) XGBoost results indicate that *Area* is the dominant factor governing improvements in IMPNSE. Rainfall characteristics (P_{sum}, P_{max}, R_{cv}) strongly influence improvements in IMP_{NSE}, IMP_{BIAS}, and IMP_{RPE}, while underlying-surface properties (P_a, *Slope*, *Clay*, *Elev*, R_{res}) also affect the extent of accuracy gains. These factors exhibit pronounced nonlinear impacts on model improvement.

(3) Optimal-resolution statistics show that fine grids (1–3 km) offer substantial benefits only when flood processes exhibit strong nonlinearity—such as in small basins, gentle slopes, or low-rainfall conditions. For larger basins or weakly nonlinear floods, performance differences across spatial scales diminish, and further refinement provides little added value.

(4) The advantages of fine grids are limited by the spatial representativeness of rainfall inputs: as rainfall-gauge density decreases, the benefits of spatial refinement weaken or even disappear. Thus, increasing spatial resolution alone cannot offset errors arising from insufficient rainfall observations.

Appendices

Appendix A

(1) Rainfall

Rainfall characteristics considered in this study include maximum rainfall intensity (P_{max}), total rainfall (P_{sum}), average rainfall intensity (P_{avg}), the rainfall spatial non-uniformity coefficient (η), and the temporal variation coefficient (R_{cv}), which together describe the magnitude, intensity, and spatiotemporal distribution of rainfall during flood events. All indicators are computed as basin-averaged values for each individual flood.

$\bar{P}(t)$epresents the basin-averaged precipitation at each time step *t*, calculated using the Thiessen polygon method. P_{max} is defined as the maximum value of $\bar{P}(t)$ during a single flood event.

$$P_{max} = \max\left(\bar{P}(t)\right) \quad \text{(A1)}$$

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

P_{sum} represents the total precipitation accumulated over all time steps, calculated as the sum of $\bar{P}(t)$ across the entire duration T of a flood event.

$$P_{sum} = \sum_{t=1}^{T} \bar{P}(t) \quad \text{(A2)}$$

P_{avg} denotes the average rainfall intensity over all wet time steps T_r (time steps with $\bar{P}(t) > 0$).

$$P_{avg} = \frac{1}{T_r} \sum_{t=1}^{T_r} \bar{P}(t) \quad \text{(A3)}$$

The rainfall spatial non-uniformity coefficient η is defined as the ratio of the basin-area-weighted total precipitation P_{sum} to the maximum cumulative precipitation among all stations $max_i(P_i)$.

$$\eta = \frac{P_{sum}}{max_i(P_i)} \quad \text{(A4)}$$

The temporal variation coefficient R_{cv} is defined as the ratio of the standard deviation of rainfall intensity to the mean rainfall intensity, reflecting the temporal concentration of rainfall. μ and σ represent the mean and standard deviation of rainfall intensity, respectively.

$$\mu = \frac{1}{T} \sum_{t=1}^{T} P_t \quad \text{(A5)}$$

$$\sigma = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T} (P_t - \mu)^2} \quad \text{(A6)}$$

$$R_{cv} = \frac{\sigma}{\mu} \quad \text{(A7)}$$

(2) Underlying Surface

The underlying surface indicators include antecedent precipitation (Pa) and soil texture components such as sand, silt, and clay fractions, which characterize the initial wetness condition and the physical properties of the catchment surface. These indicators reflect the influence of soil moisture and soil composition on runoff generation and flood response.

The antecedent precipitation Pa index quantifies the soil moisture condition prior to a rainfall event and can be used to analyze the nonlinear response of the catchment and the runoff generation process. Here, Pa_t denotes the antecedent precipitation on day t, P_t is the rainfall on the same day, K_a is the decay coefficient, and W_m represents the upper limit of antecedent precipitation.

$$Pa_t = K_a \cdot (Pa_{t-1} + P_t) \quad \text{(A8)}$$

$$Pa_t = min(Pa_t, W_m) \quad \text{(A9)}$$

Sand, *Silt*, and *Clay* represent their respective mass fractions in the soil.

$$Sand = \frac{M_{sand}}{M_{total}} \quad \text{(A10)}$$

$$Silt = \frac{M_{silt}}{M_{total}} \quad \text{(A11)}$$

$$Clay = \frac{M_{clay}}{M_{total}} \quad \text{(A12)}$$

(3) Watershed

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Watershed indicators include reservoir retention capacity (R_{res}), watershed area (*Area*), watershed shape factor (*Shape*), watershed average slope (*Slope*), and watershed average elevation (*Elev*). These indicators characterize the topographic features and storage conditions of the catchment. They reflect the watershed's water retention capacity, drainage scale, shape regularity, slope tendency, and elevation, all of which have significant impacts on runoff generation and flood response.

R_{res} represents the total reservoir retention capacity within the watershed, calculated as the sum of the capacities of all reservoirs in the catchment. Here, N is the total number of reservoirs, and R_i denotes the retention capacity of the *i-th* reservoir.

$$R_{res} = \sum_{i=1}^{N} R_i \quad \text{(A13)}$$

Area represents the total area of the watershed, calculated as the sum of the areas of all polygonal subregions within the catchment.

$$Area = \sum_{i=1}^{N} A_i \quad \text{(A14)}$$

The watershed shape factor (*Shape*) is defined as the ratio of the watershed area (*Area*) to the square of the main axis length L_{axis} of the catchment.

$$Shape = \frac{Area}{L_{axis}^2} \; Shape = \frac{Area}{L_{axis}^2} \quad \text{(A15)}$$

Watershed average slope (*Slope*) is calculated using DEM data. The gradient of each DEM grid cell within the catchment is computed, where $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ represent the rate of change of elevation in the *x* and *y* directions for each cell.

$$Slope = arctan\sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \quad \text{(A16)}$$

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Appendix B

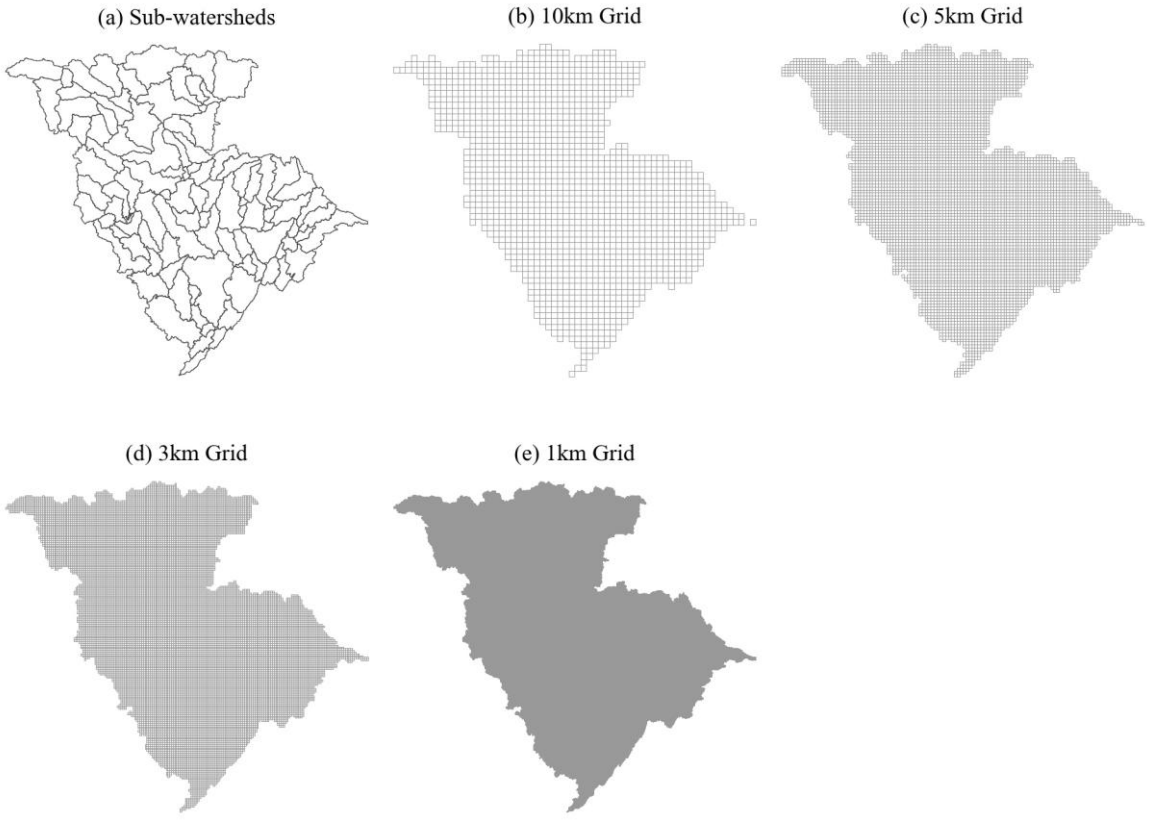


Figure B1. Sub-watershed and 10 km, 5 km, 3 km, and 1 km grid divisions of the study area

Code availability

All scripts and source code used in this study—including model implementation, data processing workflows, execution scripts, and auxiliary utilities—are publicly available on Zenodo (https://zenodo.org/records/18028215).

Data availability

The precipitation and observed streamflow data used in this study are subject to access restrictions due to confidentiality agreements with China Yangtze Power Co., Ltd. (Yichang). Access to these data may be granted upon reasonable request to China Yangtze Power Co., Ltd., and, subject to approval, the data can be provided by the co-authors.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Author contributions

Lei Ye: Data curation, Writing – review & editing, and Funding acquisition. **Xiaoyang Li:** Conceptualization, Methodology, and Writing – original draft. **Jilie Li:** Software, Visualization, and Formal Analysis. **Chi Zhang:** Writing – review and editing. **Huicheng Zhou:** Writing – review and editing.

Competing interests

The authors explicitly state that they have no financial or personal affiliations that could be construed as potential conflicts of interest regarding the content of this publication.

Acknowledgements

The authors sincerely thank China Yangtze Power Co., Ltd. (Yichang) for granting access to essential hydrological data and for their valuable support during the course of this study. This research was supported by the National Key Research and Development Program of China, the National Natural Science Foundation of China, and Dalian University of Technology, whose financial support is greatly appreciated.

Financial support

This work was supported by the National Key Research and Development Program of China (No.2023YFF0804900), the National Natural Science Foundation of China (No. 52322901)

References

Aerts JP, Hut RW, van de Giesen NC, Drost N, van Verseveld WJ, Weerts AH, Hazenberg P.: Large-sample assessment of varying spatial resolution on the streamflow estimates of the wflow_sbm hydrological model, Hydrology and Earth System Sciences, 26(16), 4407-30, https://doi.org/10.5194/hess-26-4407-2022, 2022.

Barnhart TB, Putman AL, Heldmyer AJ, Rey DM, Hammond JC, Driscoll JM, Sexstone GA.: Evaluating distributed snow model resolution and meteorology parameterizations against streamflow observations: Finer is not always better, Water Resources Research, 60(7), e2023WR035982, https://doi.org/10.1029/2023WR035982, 2024.

Chen, Y., Li, J., & Xu, H.: Improving flood forecasting capability of physically based distributed hydrological models by parameter optimization, Hydrology and Earth System Sciences, 20(1), 375-392, https://doi.org/10.5194/hess-20-375-2016, 2016.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Douinot A, Roux H, Garambois PA, Larnier K, Labat D, Dartus D.: Accounting for rainfall systematic spatial variability in flash flood forecasting, Journal of Hydrology, 541, 359-70, http://dx.doi.org/10.1016/j.jhydrol.2015.08.024, 2016.

Fischer S, Schumann AH.: Multivariate flood frequency analysis in large river basins considering tributary impacts and flood types, Water Resources Research, 57(8), e2020WR029029, https://doi.org/10.1029/2020WR029029, 2021.

Ekmekcioğlu, Ö., & Koc, K.: Explainable step-wise binary classification for the susceptibility assessment of geo-hydrological hazards, Catena, 216, 106379, https://doi.org/10.1016/j.catena.2022.106379, 2022.

Fraga I, Cea L, Puertas J.: Effect of rainfall uncertainty on the performance of physically based rainfall–runoff models, Hydrological Processes, 33(1), 160-73, https://doi.org/10.1002/hyp.13319, 2019.

Fu, X., Wang, M., Zhang, D., Chen, F., Peng, X., Wang, L., & Tan, S. K.: An XGBoost-SHAP framework for identifying key drivers of urban flooding and developing targeted mitigation strategies, Ecological Indicators, 175, 113579, https://doi.org/10.1016/j.ecolind.2025.113579, 2025.

Huang Y, Bárdossy A, Zhang K.: Sensitivity of hydrological models to temporal and spatial resolutions of rainfall data, Hydrology and Earth System Sciences, 23(6), 2647-63, https://doi.org/10.5194/hess-23-2647-2019, 2019.

Jiang R, Lu H, Yang K, Cho H, Yamazaki D.: Analysis and comparison of the flood simulations with the routing model CaMa-Flood at different spatial resolutions in the CONUS, Environmental Modelling & Software, 185, 106305, https://doi.org/10.1016/j.envsoft.2024.106305, 2025.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., & Loumagne, C.: When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events, Hydrology and Earth System Sciences, 18(2), 575-594, https://doi.org/10.5194/hess-18-575-2014, 2014.

Lovat A, Vincendon B, Ducrocq V.: Assessing the impact of resolution and soil datasets on flash-flood modelling, Hydrology and Earth System Sciences, 23(3), 1801-18, https://doi.org/10.5194/hess-23-1801-2019, 2019.

Li X, Ye L, Gu X, Chu J, Wang J, Zhang C, Zhou H.: Development of A distributed modeling framework considering spatiotemporally varying hydrological processes for Sub-Daily flood forecasting in Semi-Humid and Semi-Arid watersheds, Water Resources Management, 38(10), 3725-54, https://doi.org/10.1007/s11269-024-03837-5, 2024.

Peng S, Ding Y, Liu W, Li Z.: 1 km monthly temperature and precipitation dataset for China from 1901 to 2017, Earth System Science Data, 11(4), 1931-46, https://doi.org/10.5194/essd-11-1931-2019, 2019.

Michelon A, Benoit L, Beria H, Ceperley N, Schaefli B.: Benefits from high-density rain gauge observations for hydrological response analysis in a small alpine catchment, Hydrology and Earth System Sciences, 25(4), 2301-25, https://doi.org/10.5194/hess-25-2301-2021, 2021.

Ma K, He D, Liu S, Ji X, Li Y, Jiang H.: Novel time-lag informed deep learning framework for enhanced streamflow prediction and flood early warning in large-scale catchments, Journal of Hydrology, 631, 130841, https://doi.org/10.1016/j.jhydrol.2024.130841, 2024.

Mishra A, Mukherjee S, Merz B, Singh VP, Wright DB, Villarini G, Paul S, Kumar DN, Khedun CP, Niyogi D, Schumann G.: An overview of flood concepts, challenges, and future directions. Journal of Hydrologic Engineering, 27(6), 03122001, https://doi.org/10.1061/(ASCE)HE.1943-5584.0002164, 2022.

Mateo CM, Yamazaki D, Kim H, Champathong A, Vaze J, Oki T.: Impacts of spatial resolution and representation of flow connectivity on large-scale simulation of floods, Hydrology and Earth System Sciences, 21(10), 5143-63, https://doi.org/10.5194/hess-21-5143-2017, 2017.

Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., & Righetti, M.: Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023), Environmental Modelling & Software, 174, 105971, https://doi.org/10.1016/j.envsoft.2024.105971, 2024.

Qiao X, Nelson EJ, Ames DP, Li Z, David CH, Williams GP, Roberts W, Lozano JL, Edwards C, Souffront M, Matin MA.: A systems approach to routing global gridded runoff through local high-resolution stream networks for flood early warning systems, Environmental Modelling & Software, 120, 104501, https://doi.org/10.1016/j.envsoft.2019.104501, 2019.

Simard M, Denbina M, Marshak C, Neumann M.: A global evaluation of radar-derived digital elevation models: SRTM, NASADEM, and GLO-30, Journal of Geophysical Research: Biogeosciences, 129(11), e2023JG007672, https://doi.org/10.1029/2023JG007672, 2024.

Song, C., Fan, C., Zhu, J., Wang, J., Sheng, Y., Liu, K., ... & Ke, L.: A comprehensive geospatial database of nearly 100 000 reservoirs in China, Earth System Science Data, 14(9), 4017-4034, https://doi.org/10.5194/essd-14-4017-2022, 2022.

Shrestha P, Sulis M, Simmer C, Kollet S.: Impacts of grid resolution on surface energy fluxes simulated with an integrated surface-groundwater flow model, Hydrology and Earth System Sciences, 19(10), 4317-26, https://doi.org/10.5194/hess-19-4317-2015, 2015.

Towner J, Cloke HL, Zsoter E, Flamig Z, Hoch JM, Bazo J, Coughlan de Perez E, Stephens EM.: Assessing the performance of global hydrological models for capturing peak river flows in the Amazon basin, Hydrology and Earth System Sciences, 23(7), 3057-80, https://doi.org/10.5194/hess-23-3057-2019, 2019.

Tian J, Liu J, Wang Y, Wang W, Li C, Hu C.: A coupled atmospheric–hydrologic modeling system with variable grid sizes for rainfall–runoff simulation in semi-humid and semi-arid watersheds: how does the coupling scale affect the results?, Hydrology and Earth System Sciences, 24(8), 3933-49, https://doi.org/10.5194/hess-24-3933-2020, 2020.

Tudaji M, Nan Y, Tian F.: Assessing the value of high-resolution rainfall and streamflow data for hydrological modeling: an analysis based on 63 catchments in southeast China, Hydrology and Earth System Sciences, 29(7), 1919-37, https://doi.org/10.5194/hess-29-1919-2025, 2025.

Yu Z, Lu Q, Zhu J, Yang C, Ju Q, Yang T, Chen X, Sudicky EA.: Spatial and temporal scale effect in simulating hydrologic processes in a watershed, Journal of Hydrologic Engineering, 19(1), 99-107, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000762, 2014.

Zhu H, Chen Y.: A study of the effect of DEM spatial resolution on flood simulation in distributed hydrological modelling, Remote Sensing, 16(16), 3105, https://doi.org/10.3390/rs16163105, 2024.

https://doi.org/10.5194/egusphere-2025-6515
Preprint. Discussion started: 2 February 2026

Zhang, L., Xin, Z., Zhang, C., Song, C., & Zhou, H.: Exploring the potential of satellite precipitation after bias correction in streamflow simulation in a semi-arid watershed in northeastern China, Journal of Hydrology: Regional Studies, 43, 101192, https://doi.org/10.1016/j.ejrh.2022.101192, 2022.