

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

Lei Ye¹, Xiaoyang Li^{1,*}, Jilie Li¹, Chi Zhang¹, Huicheng Zhou¹

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

5 *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 (157,041 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 and 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 and 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

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 resolution 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 characteristics 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.

65 This study focuses on the Jialing River Basin in China (157,041 km²) 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.

70 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?

75 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,

80 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 157,041 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, 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 drainage areas (233–157,041 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.

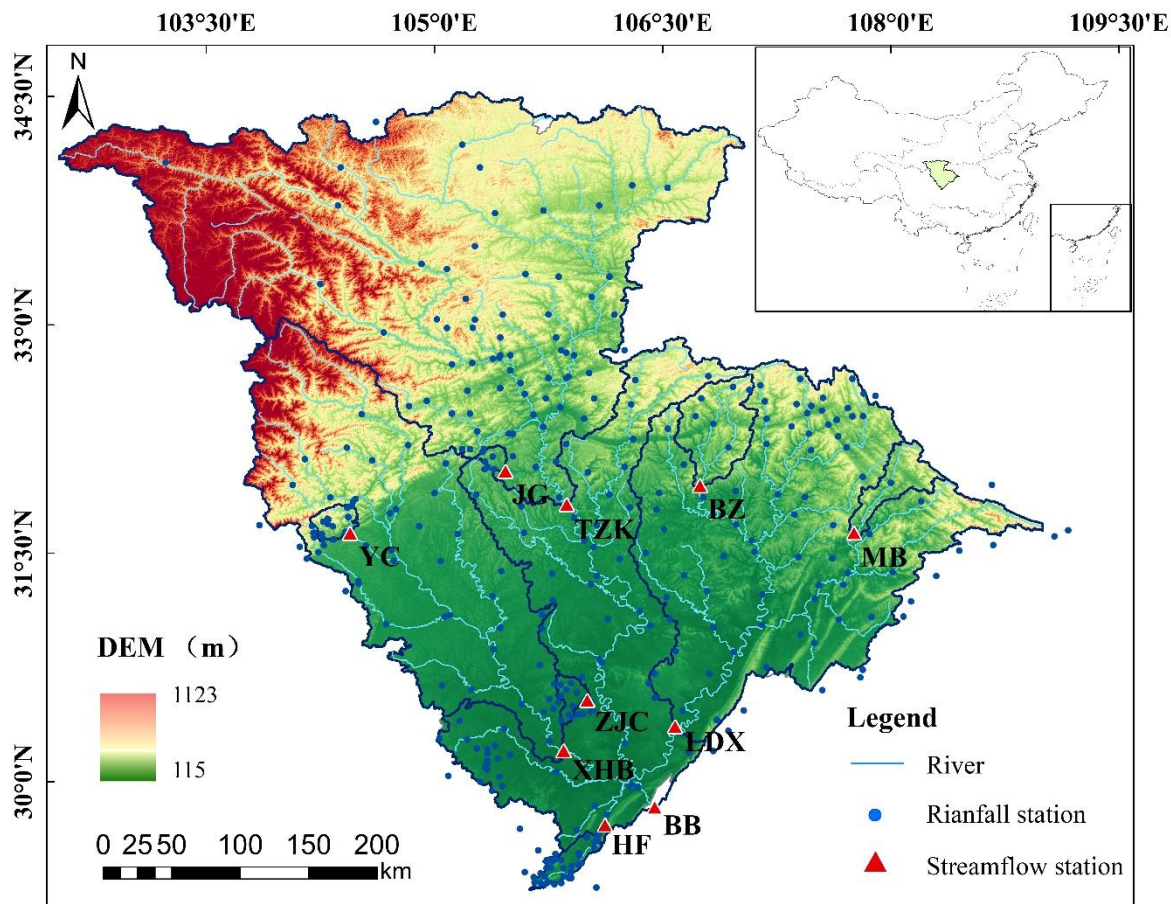


Figure 1. Location of the study area and the distribution of rainfall and streamflow stations.

95 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 to 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. The TZK station, 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.

105 A total of 282 flood events were identified based on the temporal correlation of storm–flood processes among the stations. A 30 m Digital Elevation Model (DEM) was obtained from the United States Geological Survey (<https://earthexplorer.usgs.gov/>). The 1 km soil dataset was sourced from the Food and Agriculture Organization (Batjes, 1997). 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). The watershed contains 3,806 small and medium-sized reservoirs, with a total storage capacity of 3.89 billion m³, equivalent to a runoff depth of 24.7 mm over the entire basin, mainly distributed in the middle and lower reaches.

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

Hydrological station	Abbreviation	Drainage area (km ²)	Number of rainfall stations	rainfall station density (km ² ·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
 115 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.

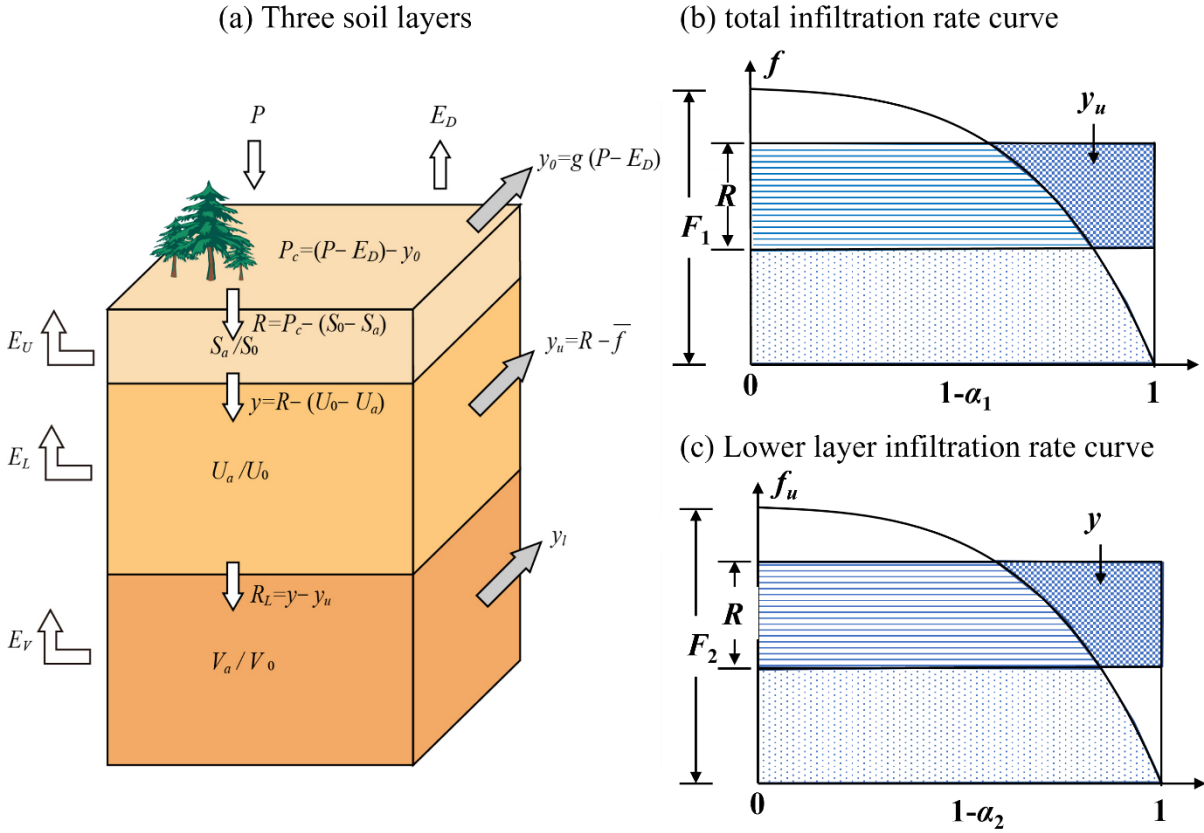
120 Fig. 2 illustrates the soil delineation and runoff generation process of the GDHF model. Soil is divided into three layers—surface, lower, and deep—while a two-layer infiltration curve represents the spatial heterogeneity of the total infiltration rate f and the lower-layer infiltration rate f_u , 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 are 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.

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).

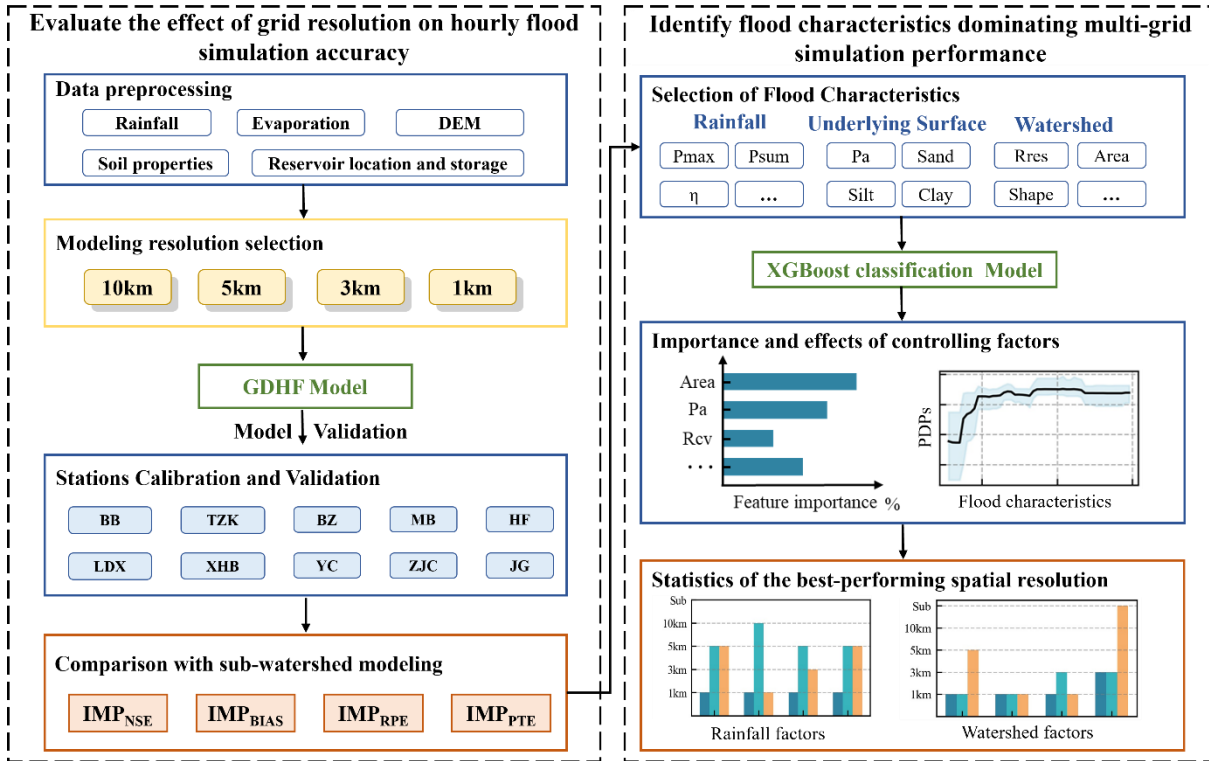


130 **Figure 2. Structural diagram of the GDHF hydrological model (Li et al., 2024).**

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 resolutions 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 resolution-sensitive controlling factors, and reveal the nonlinear effects of these factors on simulation 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.



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

2.3.1 Evaluation of GDHF Model Performance at Different Spatial Resolutions

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. The relative error of runoff (BIAS) was employed to assess the accuracy of the runoff generation process, while the relative error of flood peak (RPE) and peak timing error (PTE) were used to evaluate routing performance. The Nash–Sutcliffe efficiency coefficient (NSE) was utilized to assess the overall model performance. In Table 2, Q_{sim} and

Q_{obs} refer to the simulated and observed discharge, respectively, while T_{sp} and T_{op} denote the simulated and observed peak flow times.

155 **Table 2.** Statistical metrics used for calibration and evaluation of the GDHF model

Statistical metrics	Meaning	Equation	Perfect value
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:

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

165 $IMP_{BIAS} = |BIAS_{sub-watershed}| - Min\{|BIAS_{Grid,i}|\} \quad (4)$

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

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

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

2.3.2 Selection of Flood Characteristic Indicators

170 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. These indicators, derived from existing studies (Wang et al. 2021; Liu et al. 2022; Zhang et al. 2025), effectively represent rainfall, underlying surface conditions, and watershed characteristics. Descriptions of these indicators are provided in Table 3, with detailed calculation methods given in Appendix A.

175 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.

Table 3. Descriptions of the fourteen 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
Underlying Surface	Antecedent precipitation	P_a	mm	Reflects the antecedent soil moisture condition
	Sand	<i>Sand</i>	%	Average sand fraction of the grid
	Silt	<i>Silt</i>	%	Average silt fraction of the grid
	Clay	<i>Clay</i>	%	Average clay fraction of the grid
Watershed	Reservoir retention capacity	R_{res}	mm	The reservoir's regulation capacity for flood within the watershed
	Watershed area	<i>Area</i>	km ²	The total area encompassed by the watershed
	Watershed shape factor	<i>Shape</i>	-	Reflects the shape characteristics of the watershed
	Watershed average slope	<i>Slope</i>	degree	Indicates the variation in terrain relief
	Watershed average	<i>Elev</i>	m	Reflects the overall elevation of the

2.3.3 XGBoost classification model

185 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 resolution. XGBoost, a gradient-boosting decision tree algorithm, is widely used for hydrological applications due to its 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:

$$190 \hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (7)$$

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. Four IMP metrics and fourteen 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
195 into binary labels: label = 1 if spatial refinement yields a significant improvement, otherwise, label = 0 (Ekmekcioğlu and Koc, 2022). Because minor improvements in IMP metrics may arise from noise or random fluctuations, labeling all positive changes as 1 could lead to unstable model training and poor generalization (Siam et al., 2022). This study adopts specific thresholds ($IMP_{NSE} > 0.10$, $IMP_{BIAS} > 5\%$, $IMP_{RPE} > 5\%$, or $IMP_{PTE} > 1$ h) to identify flood events where spatial grid refinement yields significant improvements in model accuracy and performance. These thresholds are based on the common
200 ranges and practical significance of the metrics in hydrological modelling.

(b) Model training. A site-specific leave-one-station-out validation strategy was 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
205 hyperparameter tuning. This design ensures spatial independence between the training and test data, thereby effectively preventing spatial information leakage.

(c) Feature-effect interpretation. Since our study focuses on revealing the global average effects of flood characteristics, partial dependence plots (PDPs) are adopted as a global interpretation method. Compared with the SHAP method, PDPs provide a clearer and more intuitive visualization of the marginal effect of each flood characteristic on model predictions.
210 Moreover, PDPs have been widely used in previous studies to investigate the nonlinear impacts of key flood features (Wang et al., 2026; Yao et al., 2026). Therefore, PDPs were chosen 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

215 To systematically identify appropriate spatial resolutions under varying flood characteristics, this study employs three
 resolution-selection schemes based on the model performance. The purpose of setting three different schemes is to select the
 optimal spatial resolution from both accuracy and computational efficiency perspectives, as the optimal resolution is not
 unique. In the S1 scheme, the optimal spatial resolution is determined solely based on flood simulation accuracy. In contrast,
 the S2 and S3 schemes, while satisfying predefined accuracy thresholds (e.g., NSE > 0.8 for S2 and relative error ≤ 5% for
 220 S3), prioritize selecting coarser grids to improve computational efficiency. Observing the differences in the optimal spatial
 resolutions obtained from these schemes provides valuable insight for selecting appropriate spatial resolutions in large-
 watershed modeling. 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
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

225 The GDHF model 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. B1, 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.

230 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.

235 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.

240 regions.

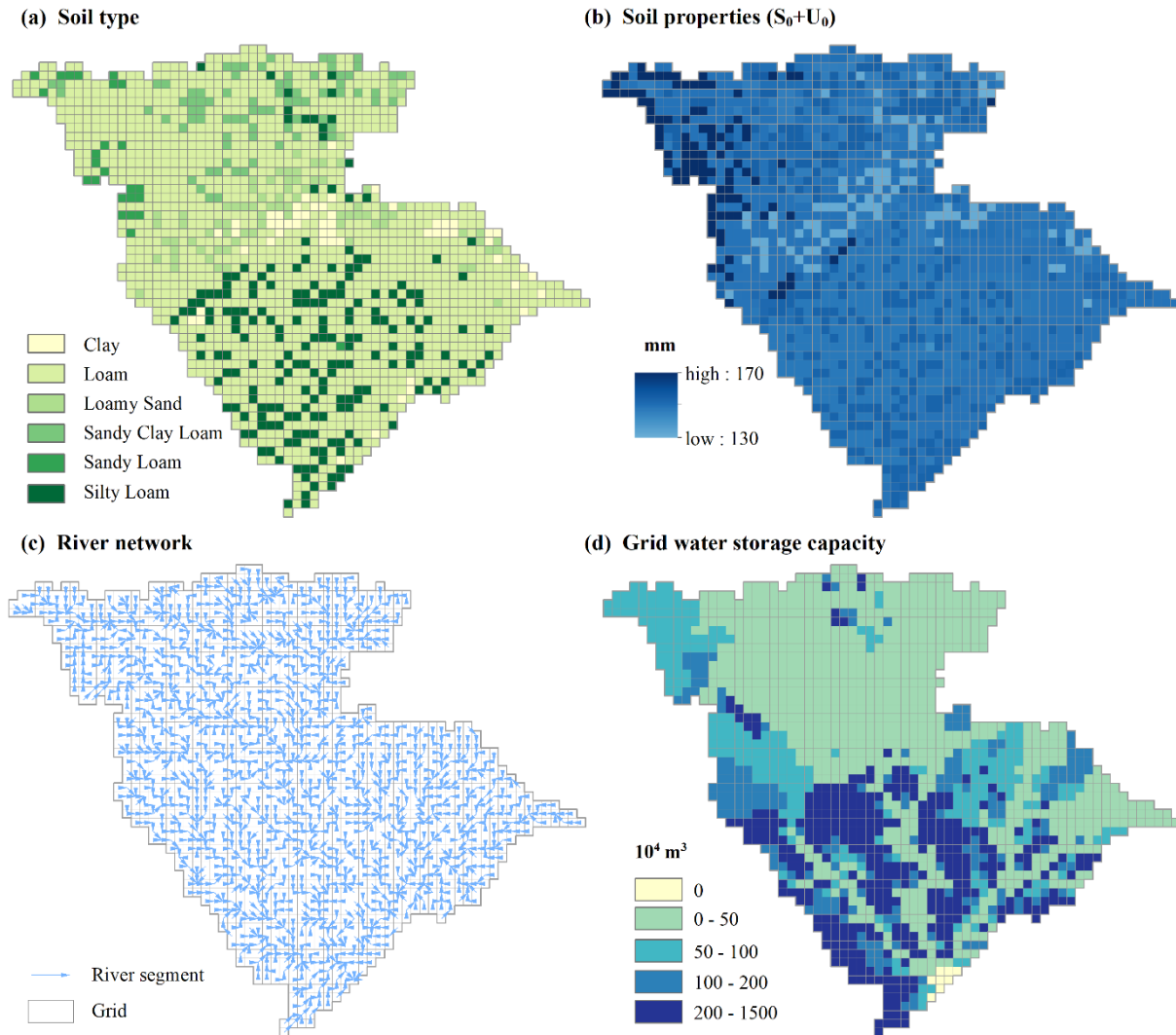


Figure 4. Spatial distribution of (a) the soil types, (b) soil properties, (c) river network, and (d) grid water storage capacity.

4 Results and Discussion

4.1 Comparison of flood simulation accuracy at different spatial resolutions

245 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 $|\text{BIAS}|$, $|\text{RPE}|$, and $|\text{PTE}|$ 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 $|\text{BIAS}|$ and $|\text{RPE}|$ remain around 10% and 15%,
250 respectively, and $|\text{PTE}|$ 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
255 drainage 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. Aerts et al. (2022) also observed this phenomenon using the daily CAMELS dataset: finer spatial resolution does not bring significant improvement at the watershed outlet. In contrast, for stations with smaller drainage areas (BZ, MB, HF, YC, ZJC, and JG), refining the spatial resolution substantially enhances the representation of runoff generation and flow routing, thereby
260 significantly improving simulation accuracy (Modi et al., 2025).

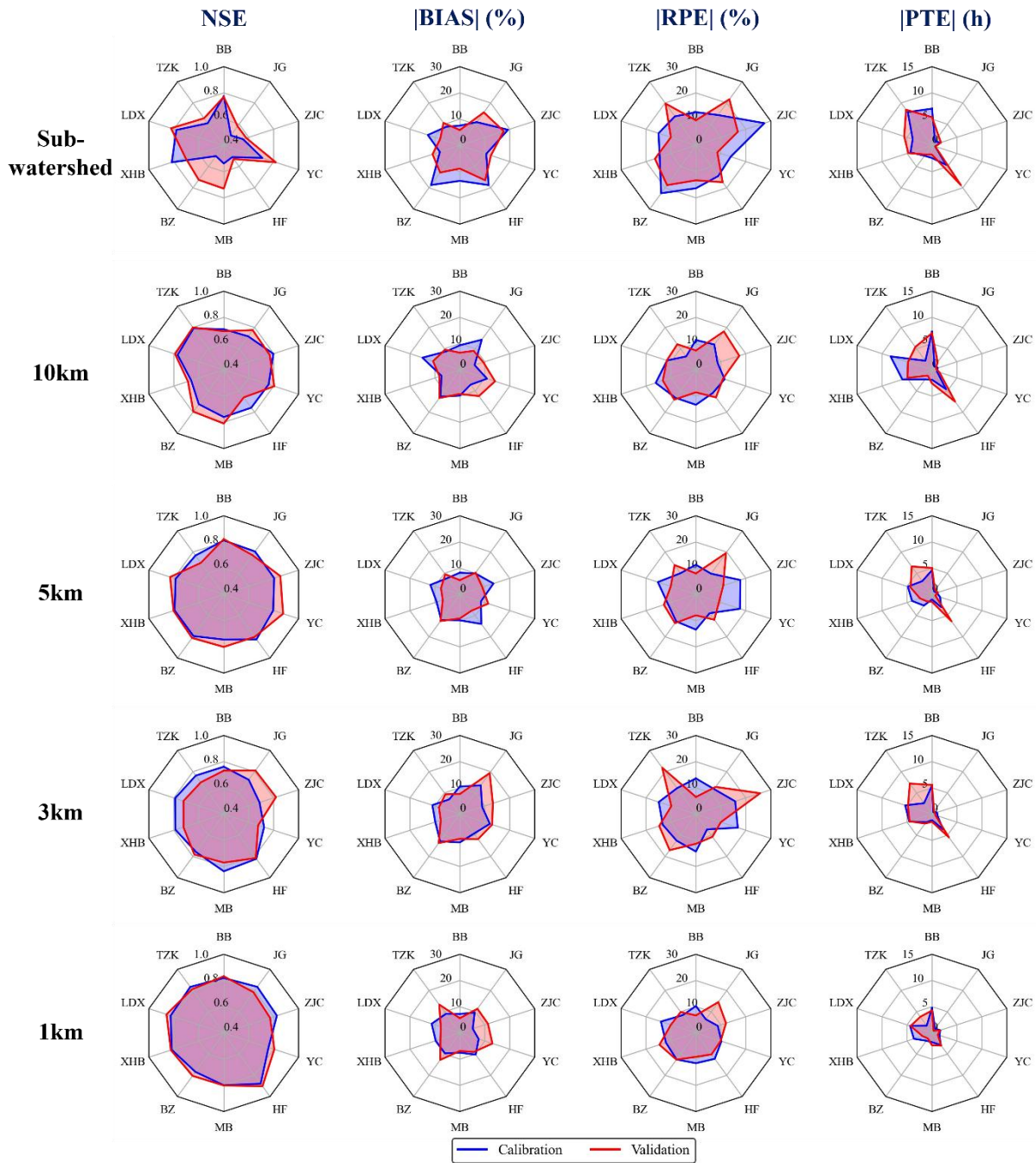
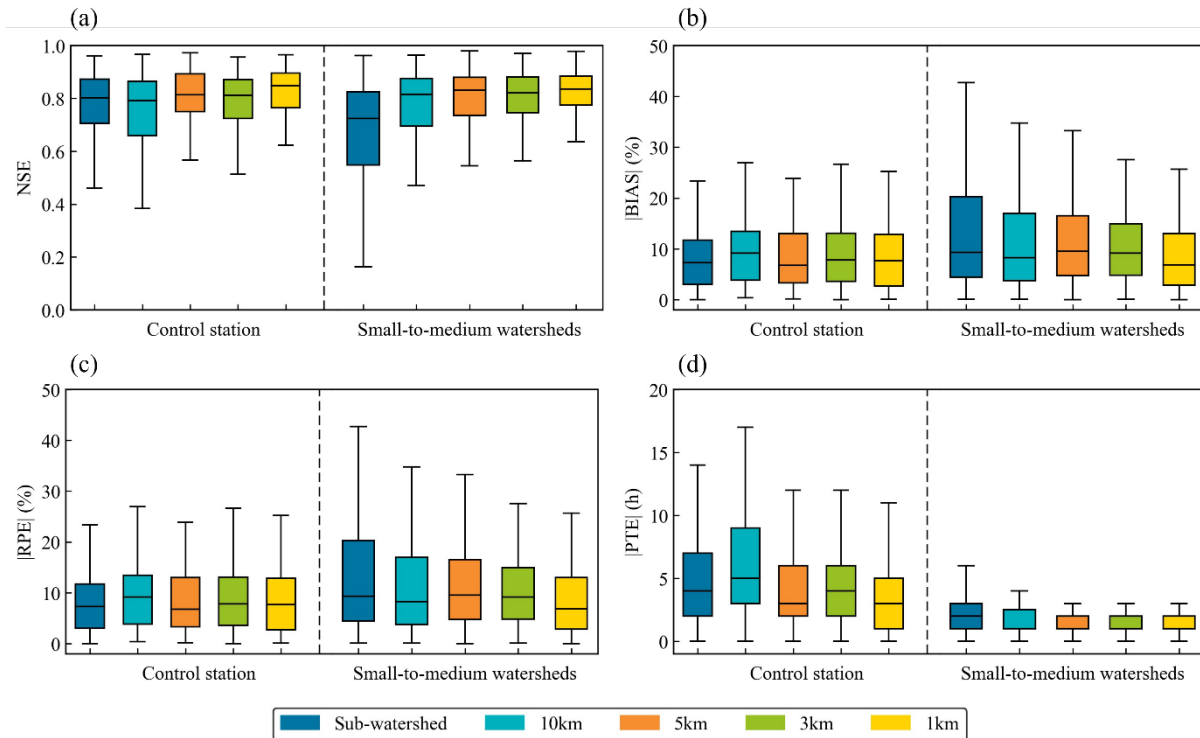


Figure 5. Evaluation metrics for simulating hourly flood events across different spatial resolutions during calibration and validation periods.

To further assess spatial resolution effects for stations of different areas, the stations were classified into two categories: 265 main-stream control stations and small-to-medium watersheds. The control stations (BB, TZK, LDX, and XHB) are located

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

Fig. 6(a) shows that for small and medium-sized 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. In large watersheds, floods originate from multiple sub-regions, and nonlinear processes become smoothed during flow aggregation, reducing the overall nonlinearity of the flood response. Small watersheds are highly sensitive to localized rainfall, steep gradients, and rapid flow routing, leading to strong nonlinear responses (Wilkinson and Bathurst, 2018). 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.



280

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

4.2 Comparison of computational time at different spatial resolutions

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.

285 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. Aerts et al. (2022) also found that computational time does not scale linearly with the number of grid cells when running models at different spatial resolutions on high-performance computing equipment. Moreover, real-time flood 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 spatial resolution to a large watershed significantly increases computational cost, making it difficult to meet the efficiency requirements of real-time flood forecasting.

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

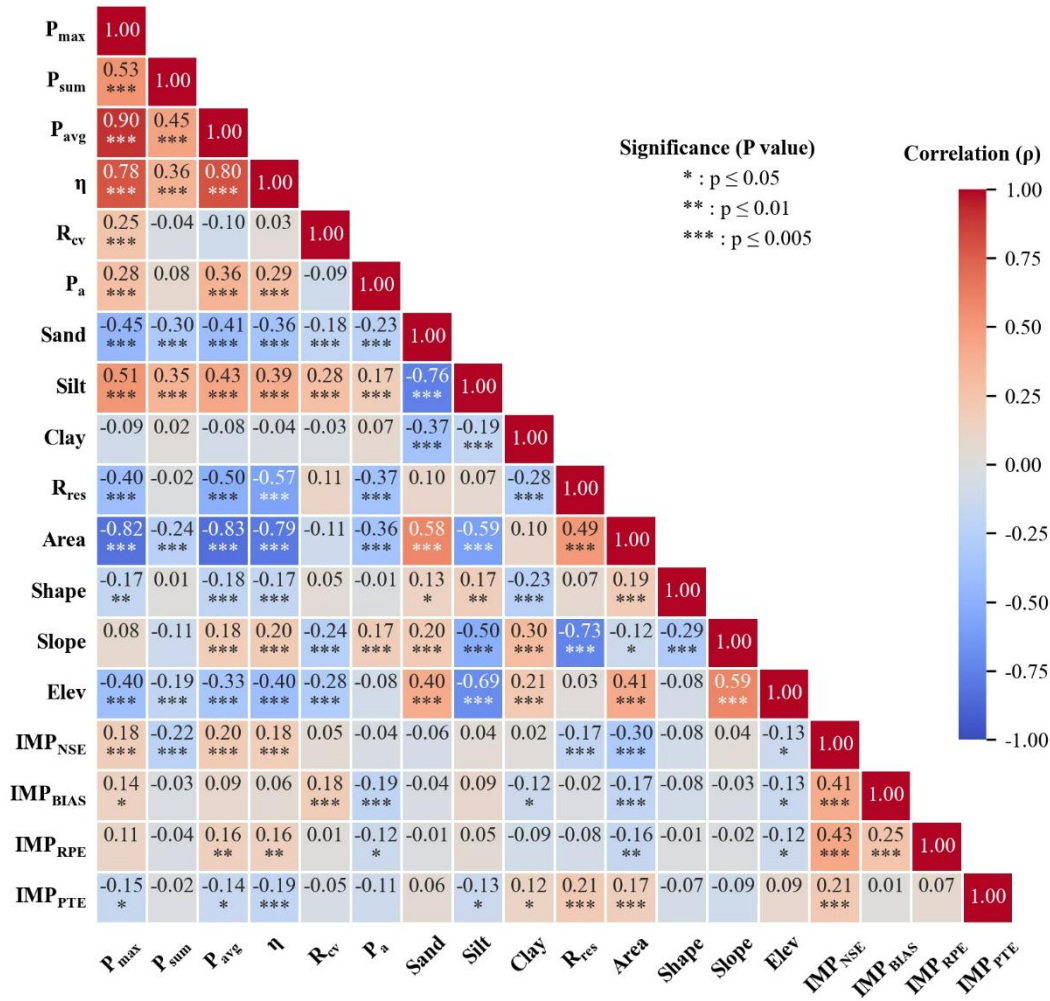
Spatial resolution	Sub-watershed	10km	5km	3km	1km
Number of cells	107	1662	6353	17283	152243
Computational time	13s	1 min 11 s	1 min 40 s	2 min 47 s	18 min 56 s

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 resolution. To ensure their robustness, we analyzed their relationships with the IMP metrics. Fig. 7 presents the Spearman correlation coefficients (ρ) and significance levels for fourteen flood characteristics and four 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 ($\rho = 0.90$), 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} ($\rho = 0.78$ and 0.80). Regarding underlying surface characteristics, $Area$ shows strong negative correlations with P_{max} , P_{avg} , and η ($\rho = -0.82$, -0.83 , and -0.79), suggesting that as $Area$ increases, rainfall intensity decreases and rainfall spatial variability increases. $Elev$ is negatively correlated with $Silt$ ($\rho = -0.69$) and positively correlated with $Slope$ ($\rho = 0.59$). R_{res} also shows a significant negative correlation with $Slope$ ($\rho = -0.73$), indicating that reservoir regulation weakens in higher-elevation watersheds.

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



315 **Figure 7. Spearman correlation analysis results between fourteen flood characteristics and four IMP metrics.**

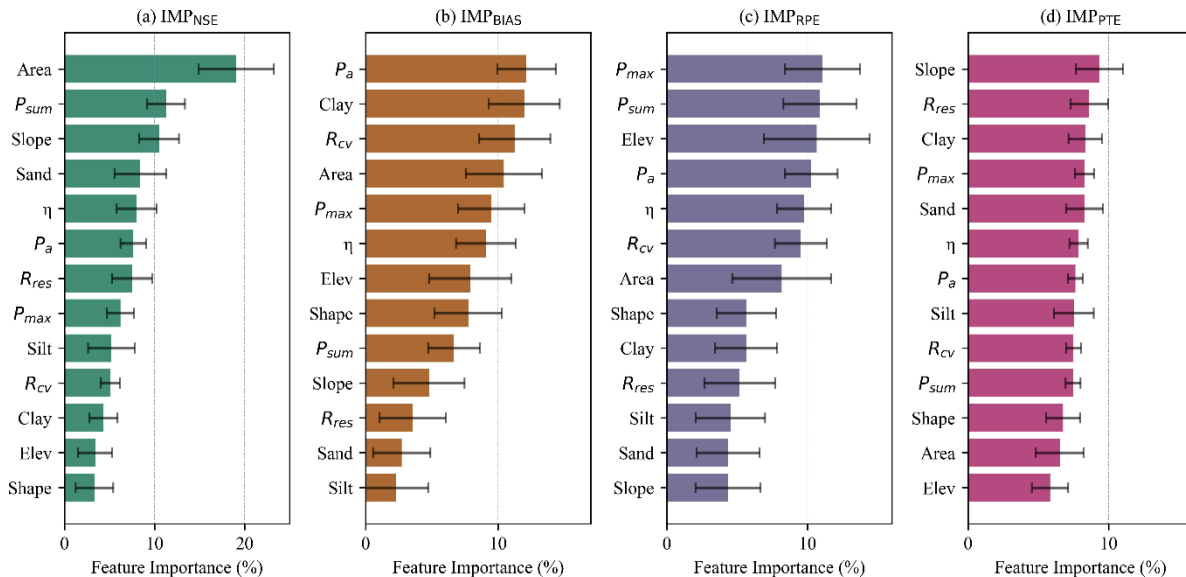
To identify how potential driving factors influence flood simulations across different spatial resolutions, this study employs an XGBoost model to identify key resolution-sensitive factors, with model configurations detailed in Section 2.3.3. Because the ρ between P_{max} and P_{avg} exceeds 0.90, indicating strong multicollinearity, P_{avg} was excluded, and P_{max} was retained as the representative rainfall intensity indicator. For all four IMP metrics, 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, with corresponding AUC of 0.78–0.98

320

(training) and 0.65–0.76 (test). These results indicate satisfactory fitting capability and acceptable generalization performance, comparable to the AUC range (0.75–0.83) reported in similar XGBoost-based classification studies (Kumar et al., 2026).

Fig. 8 presents the mean contribution rates of the driving factors and their standard deviations across the IMP metrics. For 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. Previous studies have also confirmed the importance of drainage area for the selection of spatial resolution by evaluating the improvement in accuracy achieved through spatial resolution refinement across different drainage areas (Aerts et al., 2022; Jiang et al., 2025).

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. Moreover, Luo et al. (2025) identified rainfall, elevation, and slope as key drivers of flash floods using the SHAP method, and these factors show strong consistency with the key controlling factors of spatial resolution identified in this study. Overall, the XGBoost analysis reveals that the flood characteristics controlling improvements in multi-grid simulation performance differ substantially across the various IMP evaluation metrics.

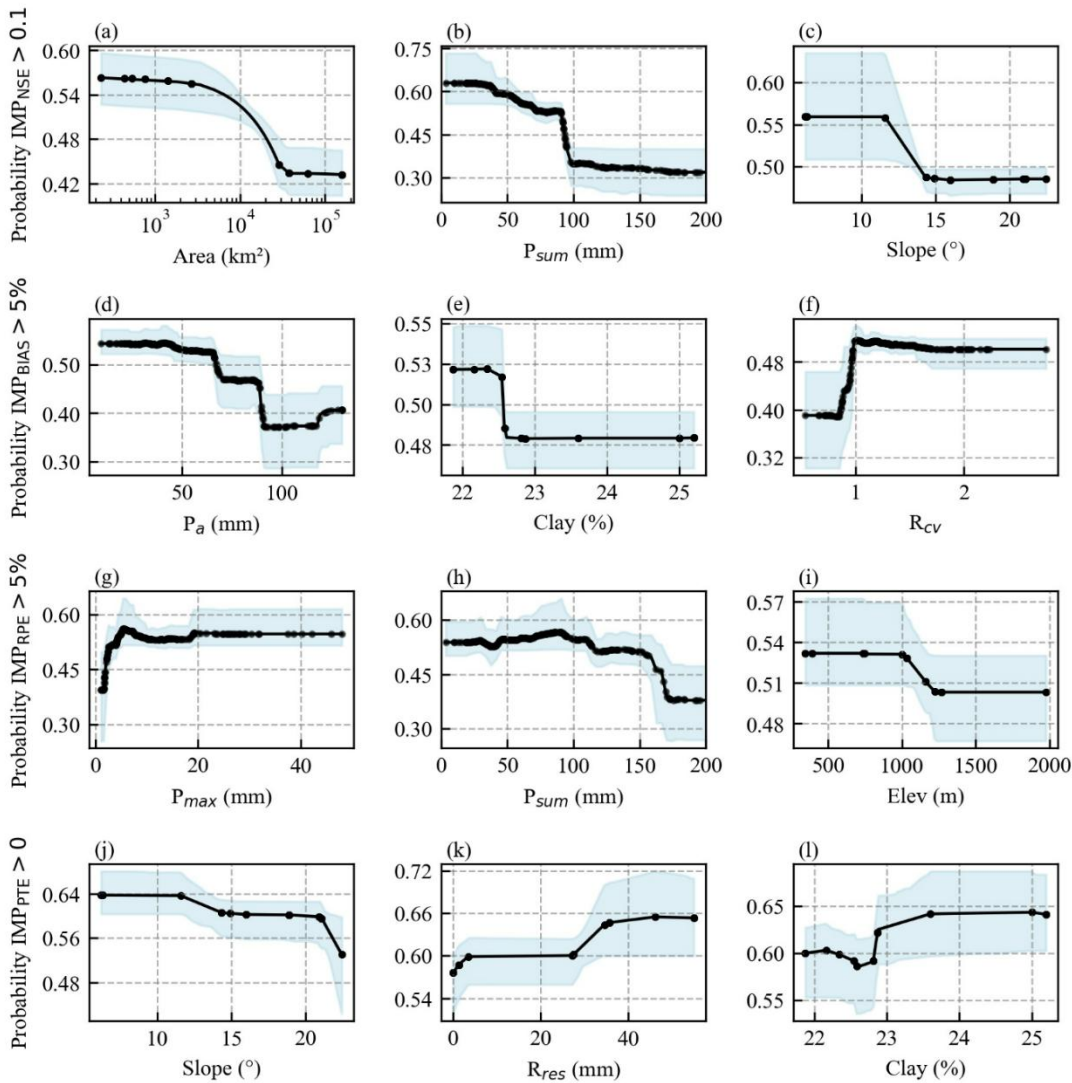


340 **Figure 8. Importance ranking and contribution rates of flood characteristics based on the XGBoost model.**

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.

- 345 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} .
- 350 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 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.
- 355 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.



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

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.

365 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 watershed area increases, the median IMP_{NSE} approaches zero, indicating that spatial refinement offers limited NSE 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 $Slope < 10^\circ$,
370 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
375 shown in Fig. 10(f), the improvement in IMP_{BIAS} is particularly significant when the rainfall temporal variability coefficient
(R_{cv}) exceeds 1.5.

Fig. 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. Fig. 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 $Slope < 10^\circ$ yields the greatest reduction in
380 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.

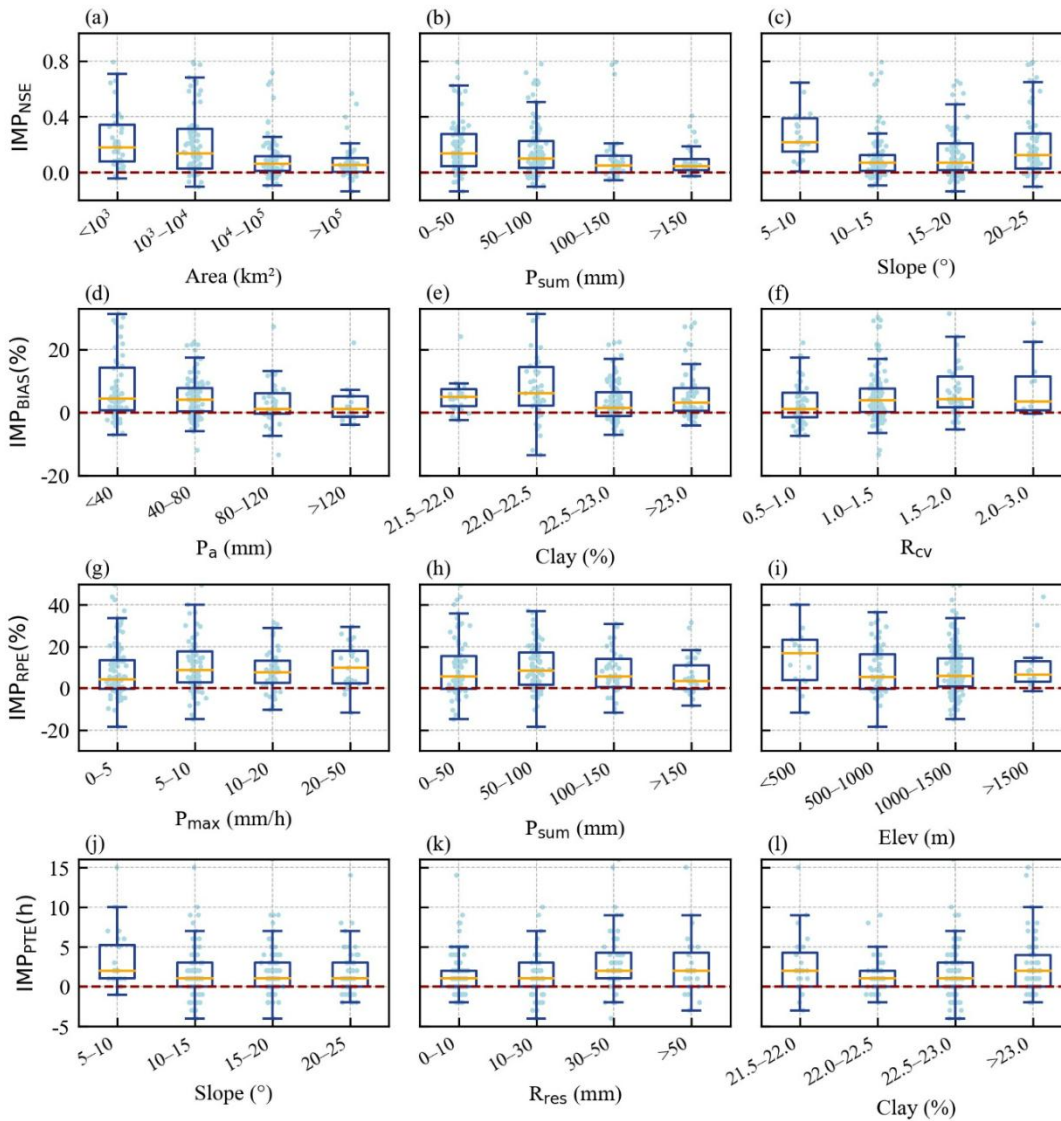


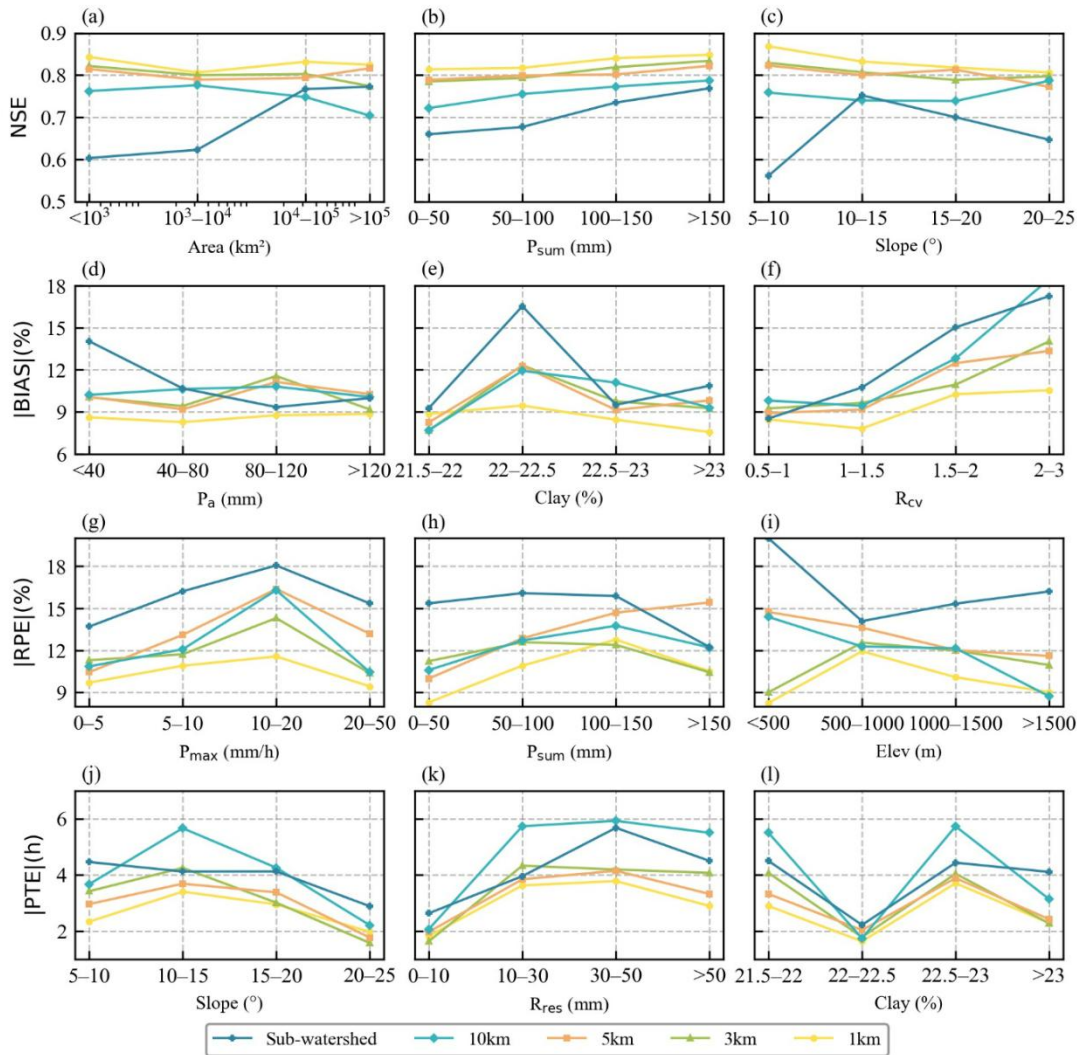
Figure 10. Boxplots of IMPNSE, IMPBIAS, IMPRPE, and IMPPTE metrics under different flood characteristics.

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 resolution under 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 ($Area < 10^4 \text{ km}^2$), gentle terrain ($Slope < 10^\circ$), or low cumulative rainfall ($P_{sum} < 100 \text{ mm}$), where the 1 km, 3 km, and 5 km grids all outperform the sub-watershed approach. In contrast, for large basins ($Area > 10^4 \text{ km}^2$), wetter antecedent soil conditions ($P_a > 80 \text{ mm}$), 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, Fig. 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.



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

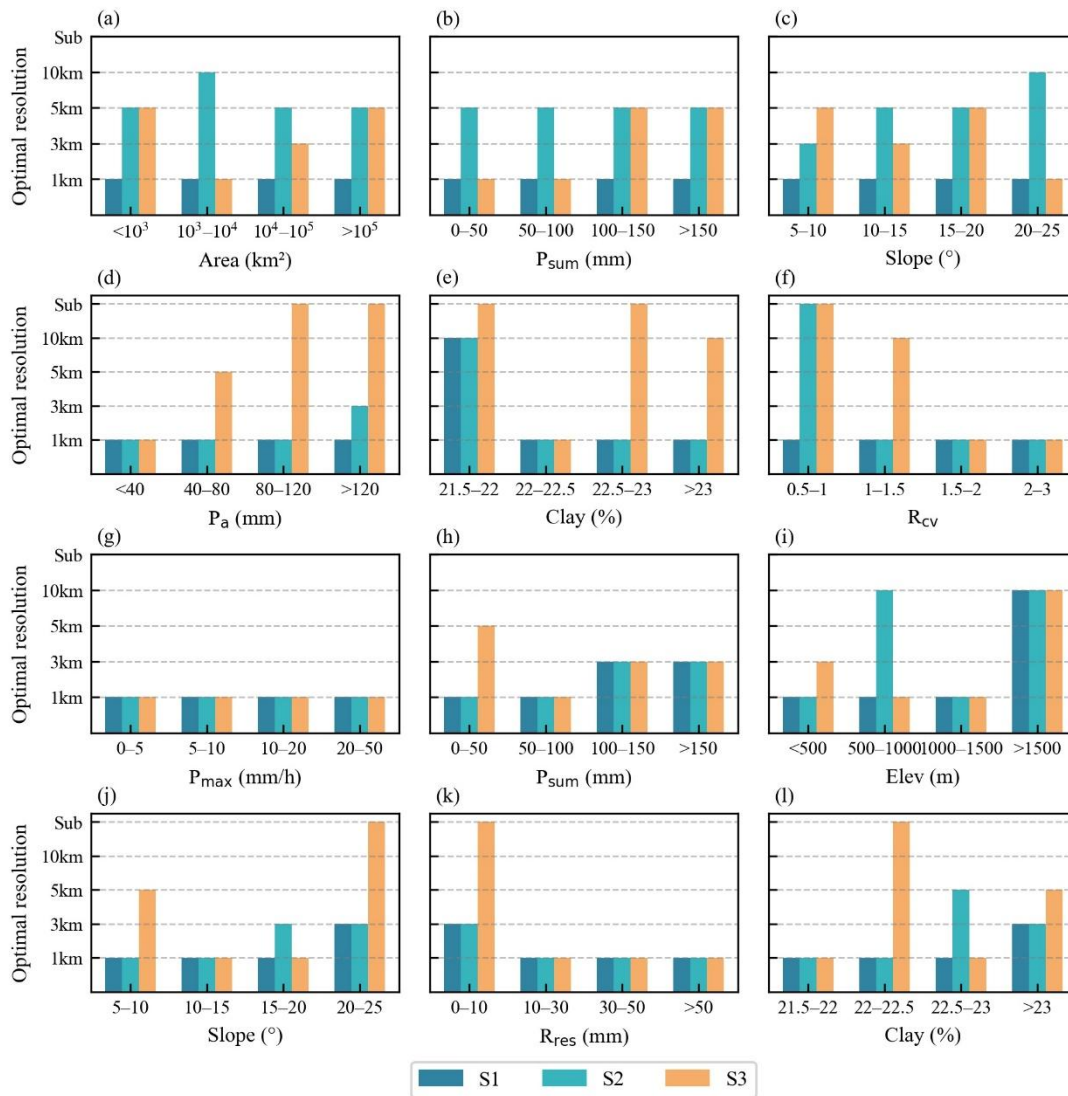
To systematically identify appropriate spatial resolutions under varying flood characteristics, this study employs three resolution-selection schemes (Section 2.3.4) based on the performance patterns observed in Fig. 11. Fig. 12 presents the appropriate spatial resolutions for different flood characteristic 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. 410 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 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 415 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 420 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 425 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 resolutions in large-watershed flood forecasting.



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

4.6 Effect of rainfall station 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 station density on flood-simulation accuracy across multiple spatial resolutions. To avoid representation issues associated with random gauge removal, the Thiessen
 435 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 station 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 station density is substantially reduced (with 50% or 75% of the gauges removed), the differences in simulation accuracy among different spatial resolutions 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. Huang et al. (2019) also found that, in watersheds with sparse gauge networks, refining spatial resolution yields model performance similar to that of the lumped model. Furthermore, Ziaee and Abedini (2023) also emphasized that rain gauge density is a key factor affecting peak flow, and insufficient monitoring station density often leads to underestimation of the peak flow. Therefore, when selecting the spatial resolution for modeling, the spatial representativeness of the rainfall data must be carefully considered.

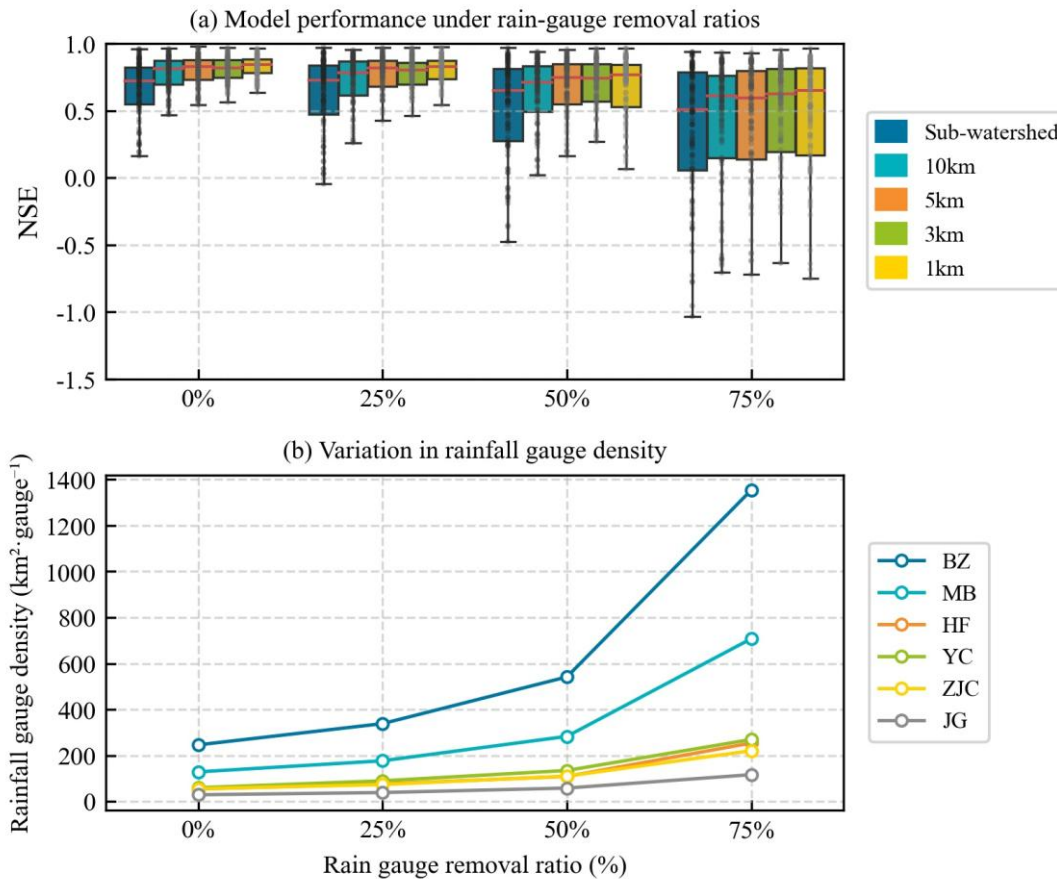
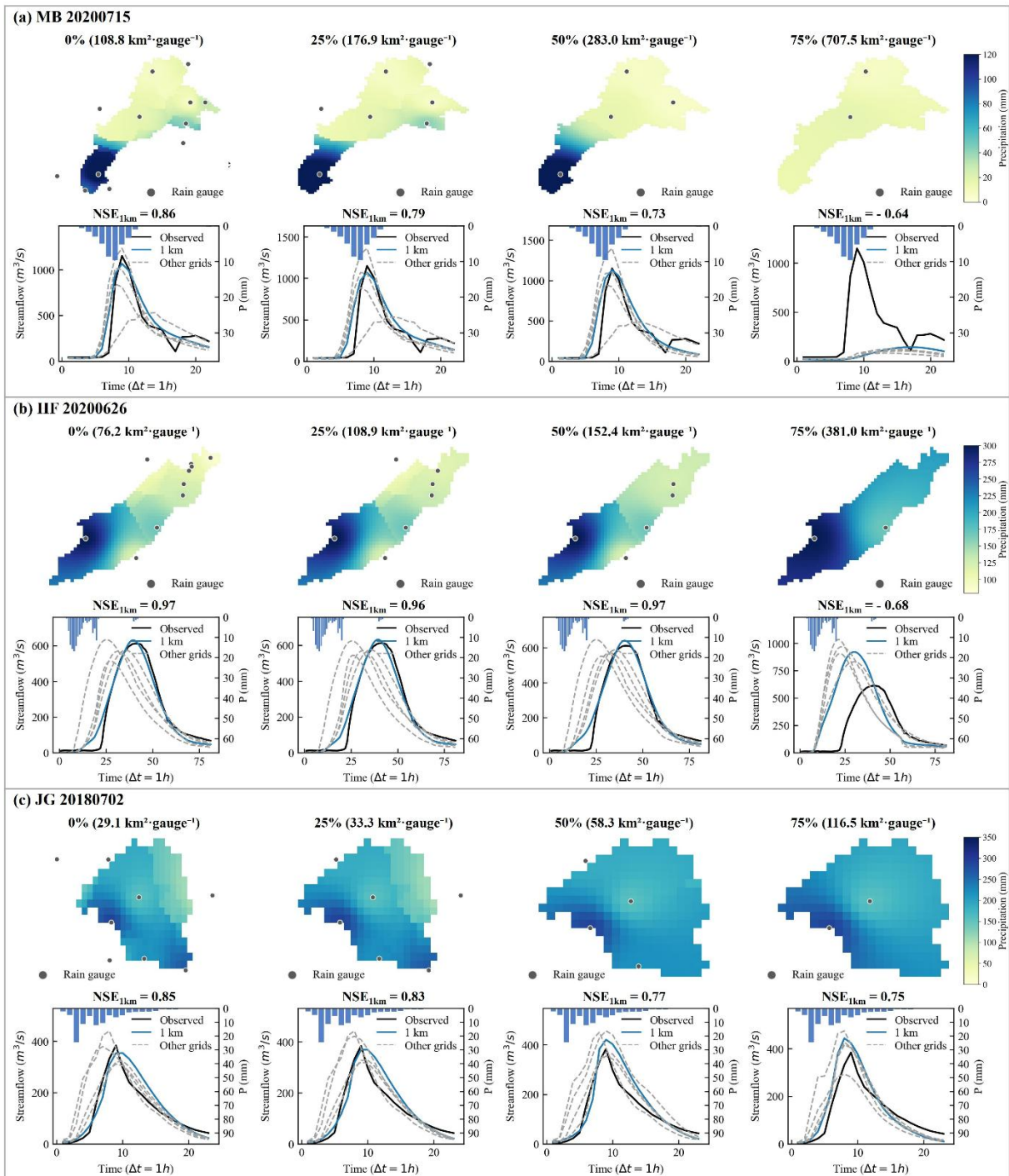


Figure 13. Flood simulation accuracy across different spatial resolutions under varying rainfall station removal ratios (0%, 25%, 50%, and 75%).

Building on this analysis, three representative flood events with notable accuracy gains from spatial-resolution refinement
455 were selected to further assess whether fine-resolution grids (1 km) remain effective under reduced rainfall station 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 $707.5 \text{ km}^2 \cdot \text{gauge}^{-1}$), the
rainfall field becomes increasingly uniform, its overall magnitude weakens, and the original storm center disappears. When
460 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 ($\text{NSE}_{1\text{km}} = 0.86, 0.79, \text{ and } 0.73$). However, with 75%
gauge removal, the representativeness of the rainfall field deteriorates sharply, leading to a substantial decline in accuracy
($\text{NSE}_{1\text{km}} = -0.64$).

Figs. 14(b) and (c) exhibit the same pattern: decreasing station density causes the rainfall field to deviate more from the true
465 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. Pan et al. (2024) demonstrated that coarser spatial resolutions of rainfall data degrade simulation accuracy.
Similarly, Michelon et al. (2021) showed in a small Alpine catchment that high-density rainfall observations are essential for
470 capturing key hydrological response characteristics, including runoff volume and peak flow timing. Therefore, when
selecting 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.



475 **Figure 14. Effects of rainfall station density on rainfall spatial distribution and flood simulation performance under different spatial resolutions.**

5 Conclusions

To investigate the influence of spatial resolution 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 station 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 substantially improve the representation of runoff generation and routing processes.

(2) XGBoost results indicate that *Area* is the dominant factor governing improvements in IMP_{NSE} . 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 resolutions 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 station density decreases, the benefits of spatial refinement weaken or even disappear. Thus, increasing spatial resolution alone cannot offset errors arising from insufficient rainfall observations.

Appendix A

(1) Rainfall

This study selected the indicators P_{max} , P_{sum} , P_{avg} , η , and R_{cv} to describe the magnitude, intensity, and spatiotemporal distribution of rainfall during flood events. Except for η , the remaining indicators are computed as watershed-averaged values for each individual flood event. The maximum rainfall intensity P_{max} is defined as the maximum value of watershed-averaged rainfall:

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

Where $\bar{P}(t)$ represents the watershed-averaged rainfall at each time step t (mm), calculated using the Thiessen polygon method.

The total rainfall P_{sum} and the average rainfall intensity P_{avg} represent the total rainfall and the average rainfall over all time steps, respectively (Zhou et al., 2021):

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

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

510 The rainfall spatial non-uniformity coefficient η and the temporal variation coefficient R_{cv} (Wang et al., 1999) are used to describe the spatiotemporal distribution characteristics of rainfall:

$$\eta = \frac{P_{avg}}{P_{max,i}} \quad (A4)$$

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

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

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

515 where $P_{max,i}$ represents the maximum point rainfall in the watershed (mm); μ and σ represent the mean and standard deviation of rainfall intensity, respectively.

(2) Underlying Surface

The underlying surface indicators include antecedent precipitation (P_a) and soil texture components (sand, silt, and clay fractions). For sand, silt, and clay fractions, the value for each grid cell is taken as the average value over the statistical grid, and the sum of the three fractions equals 1. The antecedent precipitation P_a quantifies the soil moisture condition prior to a rainfall event and can be used to analyze the watershed's nonlinear response and the runoff generation process (Schoener and Stone, 2019):

$$P_{a,t} = K_a (P_{a,t-1} + P_t) \quad (A8)$$

Where $P_{a,t}$ is the antecedent precipitation on day t , P_t is the rainfall on the same day, and K_a is the decay coefficient.

525 (3) Watershed

Watershed indicators include R_{res} , $Area$, $Shape$, $Slope$, and $Elev$. These indicators characterize the topographic features and storage conditions of the watershed (Ma et al., 2021). R_{res} represents the total reservoir retention capacity within the watershed, calculated as the sum of the capacities of all reservoirs in the watershed. $Area$ represents the total area of the watershed. The watershed shape factor ($Shape$) is defined as the ratio of the $Area$ to the square of the main axis length L_{axis} of the watershed.

$$Shape = \frac{Area}{L_{axis}^2} \quad (A9)$$

530 Watershed average slope ($Slope$) and average elevation ($Elev$) are calculated using DEM data. The gradient of each DEM grid cell within the watershed 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.

535 $Slope = \arctan \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}$ (A10)

Appendix B

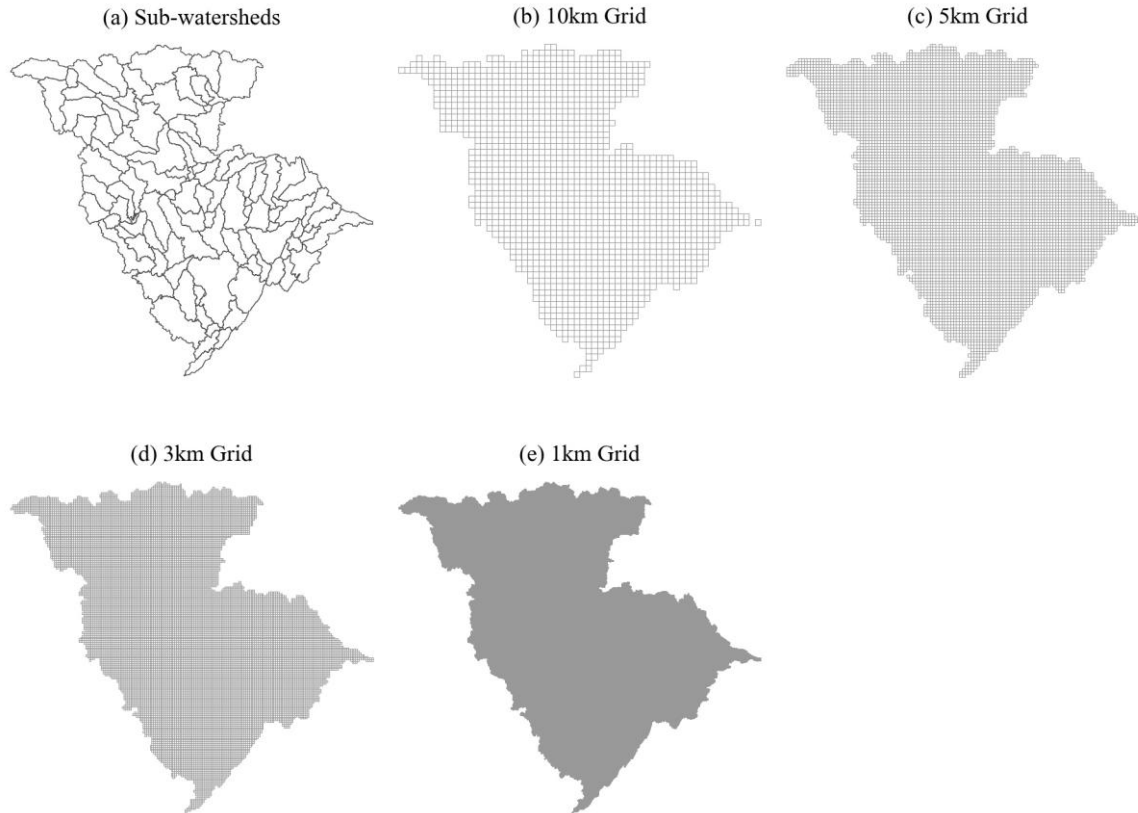


Figure B1. Sub-watershed and 10 km, 5 km, 3 km, and 1 km grid divisions of the study area

Code availability

540 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
545 China Yangtze Power Co., Ltd., and, subject to approval, the data can be provided by the co-authors.

Author contributions

Lei Ye: Data curation, Writing – review and editing, and Funding acquisition. **Xiaoyang Li:** Conceptualization, Methodology, and Writing – original draft. **Jilie Li:** Software, Visualization, and Formal Analysis. **Chi Zhang:** Writing –
550 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

555 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

560 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, J. P., Hut, R. W., van de Giesen, N. C., Drost, N., van Verseveld, W. J., Weerts, A. H., and Hazenberg, P.: Large-sample assessment of varying spatial resolution on the streamflow estimates of the wflow_sbm hydrological model, Hydrol.
565 Earth Syst. Sci., 26, 4407–4430, <https://doi.org/10.5194/hess-26-4407-2022>, 2022.

- Barnhart, T. B., Putman, A. L., Heldmyer, A. J., Rey, D. M., Hammond, J. C., Driscoll, J. M., and Sexstone, G. A.: Evaluating distributed snow model resolution and meteorology parameterizations against streamflow observations: Finer is not always better, *Water Resour. Res.*, 60, e2023WR035982, <https://doi.org/10.1029/2023WR035982>, 2024.
- Batjes, N. H.: A world dataset of derived soil properties by FAO–UNESCO soil unit for global modelling. *Soil Use Manage.*, 13, 916, <https://doi.org/10.1111/j.1475-2743.1997.tb00550.x>, 1997.
- Douinot, A., Roux, H., Garambois, P. A., Larnier, K., Labat, D., and Dartus, D.: Accounting for rainfall systematic spatial variability in flash flood forecasting, *J. Hydrol.*, 541, 359–370, <http://dx.doi.org/10.1016/j.jhydrol.2015.08.024>, 2016.
- Fischer, S., and Schumann, A. H.: Multivariate flood frequency analysis in large river basins considering tributary impacts and flood types, *Water Resour. Res.*, 57, e2020WR029029, <https://doi.org/10.1029/2020WR029029>, 2021.
- 575 Ekmekcioğlu, Ö., and Koc, K.: Explainable step-wise binary classification for the susceptibility assessment of hydrological hazards, *Catena*, 216, 106379, <https://doi.org/10.1016/j.catena.2022.106379>, 2022.
- Fraga, I., Cea, L., and Puertas, J.: Effect of rainfall uncertainty on the performance of physically based rainfall–runoff models, *Hydrol. Process.*, 33, 160–173, <https://doi.org/10.1002/hyp.13319>, 2019.
- Fu, X., Wang, M., Zhang, D., Chen, F., Peng, X., Wang, L., and Tan, S. K.: An XGBoost-SHAP framework for identifying 580 key drivers of urban flooding and developing targeted mitigation strategies, *Ecol. Indic.*, 175, 113579, <https://doi.org/10.1016/j.ecolind.2025.113579>, 2025.
- Huang, Y., Bárdossy, A., and Zhang, K.: Sensitivity of hydrological models to temporal and spatial resolutions of rainfall data, *Hydrol. Earth Syst. Sci.*, 23, 2647–2663, <https://doi.org/10.5194/hess-23-2647-2019>, 2019.
- Jiang, R., Lu, H., Yang, K., Cho, H., and Yamazaki, D.: Analysis and comparison of the flood simulations with the routing 585 model CaMa-Flood at different spatial resolutions in the CONUS, *Environ. Model. Softw.*, 185, 106305, <https://doi.org/10.1016/j.envsoft.2024.106305>, 2025.
- Kumar, A., Pandey, G., and Kale, R. V.: Ensemble machine learning and deep learning framework for flood susceptibility mapping in the transboundary Rapti River Basin. *Environ. Earth Sci.*, 85, 168, <https://doi.org/10.1007/s12665-026-12912-6>, 2026.
- 590 Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., and Loumagne, C.: When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events, *Hydrol. Earth Syst. Sci.*, 18, 575–594, <https://doi.org/10.5194/hess-18-575-2014>, 2014.
- Lovat, A., Vincendon, B., and Ducrocq, V.: Assessing the impact of resolution and soil datasets on flash-flood modelling, *Hydrol. Earth Syst. Sci.*, 23, 1801–1818, <https://doi.org/10.5194/hess-23-1801-2019>, 2019.
- 595 Li, X., Ye, L., Gu, X., Chu, J., Wang, J., Zhang, C., and 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 Resour. Manag.*, 38, 3725–3754, <https://doi.org/10.1007/s11269-024-03837-5>, 2024.
- Liu, Y., Li, Z., Liu, Z., and Luo, Y.: Impact of rainfall spatiotemporal variability and model structures on flood simulation in semi-arid regions, *Stoch. Environ. Res. Risk Assess.*, 36, 785–809, <https://doi.org/10.1007/s00477-021-02050-9>, 2022.

- 600 Luo, L., Wang, Y., Li, Q., Li, M., Wang, J., Zhao, G., and Ma, M.: Exploration of the spatiotemporal characteristics and triggering factors of flash flood in China, *Ecol. Indic.*, 176, 113698, <https://doi.org/10.1016/j.ecolind.2025.113698>, 2025.
- Michelon, A., Benoit, L., Beria, H., Ceperley, N., and Schaeffli, B.: Benefits from high-density rain gauge observations for hydrological response analysis in a small alpine catchment, *Hydrol. Earth Syst. Sci.*, 25, 2301–2325, <https://doi.org/10.5194/hess-25-2301-2021>, 2021.
- 605 Ma, W., Ishitsuka, Y., Takeshima, A., Hibino, K., Yamazaki, D., Yamamoto, K., Kachi, M., Oki, R., Oki, T., and Yoshimura, K.: Applicability of a nationwide flood forecasting system for Typhoon Hagibis 2019, *Sci. Rep.*, 11, 10213, <https://doi.org/10.1038/s41598-021-89522-8>, 2021.
- Ma, K., He, D., Liu, S., Ji, X., Li, Y., and Jiang, H.: Novel time-lag informed deep learning framework for enhanced streamflow prediction and flood early warning in large-scale catchments, *J. Hydrol.*, 631, 130841, <https://doi.org/10.1016/j.jhydrol.2024.130841>, 2024.
- 610 Mateo, C. M. R., Yamazaki, D., Kim, H., Champathong, A., Vaze, J., and Oki, T.: Impacts of spatial resolution and representation of flow connectivity on large-scale simulation of floods, *Hydrol. Earth Syst. Sci.*, 21, 5143–25163, <https://doi.org/10.5194/hess-21-5143-2017>, 2017.
- Modi, P., Yamazaki, D., Hirabayashi, Y., Revel, M., and Zhou, X.: How Spatial Resolutions Impact the Large-Scale River Hydrodynamic Model Simulations: Analysis Focuses on Model Physics, *J. Adv. Model. Earth Syst.*, 17, e2025MS004961, <https://doi.org/10.1029/2025MS004961>, 2025.
- 615 Niazkar, M., Menapace, A., Brentan, B., Piraei, R., Jimenez, D., Dhawan, P., and Righetti, M.: Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023), *Environ. Model. Softw.*, 174, 105971, <https://doi.org/10.1016/j.envsoft.2024.105971>, 2024.
- 620 Pan, X., Hou, J., Wang, T., Li, X., Jing, J., Chen, G., Qiao, J., and Guo, Q.: Study on the Influence of Temporal and Spatial Resolution of Rainfall Data on Watershed Flood Simulation Performance, *Water Resour. Manag.*, 38, 2647–2668, <https://doi.org/10.1007/s11269-023-03661-3>, 2024.
- Peng, S., Ding, Y., Liu, W., and Li, Z.: 1 km monthly temperature and precipitation dataset for China from 1901 to 2017, *Earth Syst. Sci. Data*, 11, 1931–1946, <https://doi.org/10.5194/essd-11-1931-2019>, 2019.
- 625 Qiao, X., Nelson, E. J., Ames, D. P., Li, Z., David, C. H., Williams, G. P., Roberts, W., Sánchez Lozano, J. L., Edwards, C., Souffront, M., and Matin, M. A.: A systems approach to routing global gridded runoff through local high-resolution stream networks for flood early warning systems, *Environ. Model. Softw.*, 120, 104501, <https://doi.org/10.1016/j.envsoft.2019.104501>, 2019.
- Schoener, G., and Stone, M. C.: Impact of antecedent soil moisture on runoff from a semiarid catchment, *J. Hydrol.*, 569, 627–636, <https://doi.org/10.1016/j.jhydrol.2018.12.025>, 2019.
- 630 Siam, Z. S., Hasan, R. T., and Rahman, R. M.: Effects of Label Noise on Regression Performances and Model Complexities for Hybridized Machine Learning Based Spatial Flood Susceptibility Modelling, *Cybern. Syst.*, 53, 362–379, <https://doi.org/10.1080/01969722.2021.1988446>, 2022.

- Simard, M., Denbina, M., Marshak, C., and Neumann, M.: A global evaluation of radar-derived digital elevation models: SRTM, NASADEM, and GLO-30, *J. Geophys. Res. Biogeosci.*, 129, e2023JG007672, <https://doi.org/10.1029/2023JG007672>, 2024.
- Song, C., Fan, C., Zhu, J., Wang, J., Sheng, Y., Liu, K., Chen, T., Zhan, P., Luo, S., Yuan, C., and Ke, L.: A comprehensive geospatial database of nearly 100 000 reservoirs in China, *Earth Syst. Sci. Data*, 14, 4017–4034, <https://doi.org/10.5194/essd-14-4017-2022>, 2022.
- Shrestha, P., Sulis, M., Simmer, C., and Kollet, S.: Impacts of grid resolution on surface energy fluxes simulated with an integrated surface-groundwater flow model, *Hydrol. Earth Syst. Sci.*, 19, 4317–4326, <https://doi.org/10.5194/hess-19-4317-2015>, 2015.
- Towner, J., Cloke, H. L., Zsoter, E., Flamig, Z., Hoch, J. M., Bazo, J., Coughlan de Perez, E., and Stephens, E. M.: Assessing the performance of global hydrological models for capturing peak river flows in the Amazon basin, *Hydrol. Earth Syst. Sci.*, 23, 3057–3080, <https://doi.org/10.5194/hess-23-3057-2019>, 2019.
- Tian, J., Liu, J., Wang, Y., Wang, W., Li, C., and 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?, *Hydrol. Earth Syst. Sci.*, 24, 3933–3949, <https://doi.org/10.5194/hess-24-3933-2020>, 2020.
- Tudaji, M., Nan, Y., and Tian, F.: Assessing the value of high-resolution rainfall and streamflow data for hydrological modeling: an analysis based on 63 catchments in southeast China, *Hydrol. Earth Syst. Sci.*, 29, 1919–1937, <https://doi.org/10.5194/hess-29-1919-2025>, 2025.
- Wang, Q., Xu, Y., Cai, X., Tang, J., and Yang, L.: Role of underlying surface, rainstorm and antecedent wetness condition on flood responses in small and medium sized watersheds in the Yangtze River Delta region, China, *CATENA*, 206, 105489, <https://doi.org/10.1016/j.catena.2021.105489>, 2021.
- Wang, W. Z., Jiao, J. Y., and Hao, X. P.: Nonuniformity of spatial distribution of rainfall and relationship between point rainfall and areal rainfall of different patterns of rainstorm on the Loess Plateau, *Adv Water Sci (in Chinese)*, 10, 165–169, 1999.
- Wang, T., Fu, Z., and Luo, M.: Deep learning for decoding climate–urbanization synergies: flood susceptibility forecasting under SSP-RCP scenarios in Beijing, China, *J. Hydrol.*, 672, 135364, <https://doi.org/10.1016/j.jhydrol.2026.135364>, 2026.
- Wilkinson, M. E., and Bathurst, J. C.: A multi-scale nested experiment for understanding flood wave generation across four orders of magnitude of catchment area, *Hydrol. Res.*, 49, 597–615, <https://doi.org/10.2166/nh.2017.070>, 2018.
- Yao, Y., Fu, G., and Webber, J.: How can hydrological connectivity inform catchment scale stormwater flood management?, *Water Res.*, 297, 125767, <https://doi.org/10.1016/j.watres.2026.125767>, 2026.
- Yu, Z., Lu, Q., Zhu, J., Yang, C., Ju, Q., Yang, T., Chen, X., and Sudicky, E. A.: Spatial and temporal scale effect in simulating hydrologic processes in a watershed, *J. Hydrol. Eng.*, 19, 99–107, [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000762](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000762), 2014.

- Zhu, H., and Chen, Y.: A study of the effect of DEM spatial resolution on flood simulation in distributed hydrological modelling, *Remote Sens.*, 16, 3105, <https://doi.org/10.3390/rs16163105>, 2024.
- 670 Zhang, J., Wang, R., Li, X., Li, Z., Wang, Y., and Zhang, X.: A study on the effect of spatially variation rainfall on urban flooding, *Geomatics Nat. Hazards Risk*, 16, <https://doi.org/10.1080/19475705.2025.2548912>, 2025.
- Zhang, L., Xin, Z., Zhang, C., Song, C., and Zhou, H.: Exploring the potential of satellite precipitation after bias correction in streamflow simulation in a semi-arid watershed in northeastern China, *J. Hydrol. Reg. Stud.*, 43, 101192, <https://doi.org/10.1016/j.ejrh.2022.101192>, 2022.
- 675 Zhou, Z., Smith, J. A., Baeck, M. L., Wright, D. B., Smith, B. K., and Liu, S.: The impact of the spatiotemporal structure of rainfall on flood frequency over a small urban watershed: an approach coupling stochastic storm transposition and hydrologic modelling, *Hydrol. Earth Syst. Sci.*, 25, 4701–4717, <https://doi.org/10.5194/hess-25-4701-2021>, 2021.
- Ziaee, P., and Abedini, M. J.: Investigating the Effect of Spatial and Temporal Variabilities of Rainfall on Catchment Response, *Water Resour. Manag.*, 37, 5343–5366, 2023.