International Journal of Applied Earth Observation and Geoinformation 154 (2026) 105590

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Enhanced early warning of extreme river discharge events in the Yangtze River Basin using atmospheric circulation signals

Xiaoke Xu [a,b], Yong Wang [b], Anning Huang [c,d], Yinghong Jing [b], Chunlei Gu [c], Xiaojun She [b], Lifu Zhang [e], Yao Li [b,e,*]

[a] School of Computer, Chengdu University of Information Technology, Chengdu 610225, China
[b] School of Geographical Sciences, Southwest University, Chongqing 400715, China
[c] School of Atmospheric Sciences, Nanjing University, Nanjing 210023, China
[d] Plateau Atmosphere and Environment Key Laboratory of Sichuan Province, College of Atmospheric Sciences, Chengdu University of Information Technology, Chengdu, China
[e] State Key Laboratory of Remote Sensing Science and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences. Beijing 100101, China

ARTICLE INFO

Keywords:
Atmospheric circulation signals
Extreme river discharge events
Precipitation
Runoff
Yangtze River Basin

ABSTRACT

Early warning of extreme river discharge events (ERDEs) is crucial for water resources management and disaster reduction in the Yangtze River Basin (YRB), yet substantial uncertainty remains in their prediction. This study proposes a novel strategy to enhance ERDE early warning by detecting its associated atmospheric circulation signals. We develop DetRF, a machine learning model that integrates anomaly detection with a balanced random forest ensemble. DetRF was trained and validated using ERA5 reanalysis data from 2000 to 2019 and independently applied to 2020–2024, with its early warning performance evaluated against satellite-based precipitation and reanalysis river discharge data. Across eight subregions of the YRB, DetRF effectively detects ERDE-related atmospheric circulation signals, achieving mean recall, precision, and F1 score values of 0.756, 0.506, and 0.576, respectively, and consistently outperforming benchmark machine learning models. The detected signals accurately capture key synoptic features associated with hydrological extremes, including the positions and intensities of troughs, vortices, and cyclones, as well as moisture transport. More than 99% of the detected signals are followed by precipitation within five days, while persistent precipitation occurs in over 67% of cases across both local and upstream regions. More importantly, an average of 93.8% of DetRF-detected circulation signals are followed by at least one ERDE within the subsequent five days. These findings establish a direct link between specific atmospheric circulation signals and subsequent hydrological extremes, offering a new pathway for extending ERDE early warning beyond conventional hydrological prediction. The proposed framework is particularly promising for regions where hydrological observations are sparse.

1. Introduction

The Yangtze River Basin (YRB) sustains one-third of China's population, making water security a national priority (Wei et al., 2020; Jia et al., 2022). Although flood disasters historically account for most disaster-related losses, the proactive management of extreme river discharge events (ERDEs)—defined as daily river discharge exceeding the 95th percentile of the local historical record—has become increasingly urgent (Yang et al., 2021; Yang et al., 2022; Xu et al., 2026). Under a warming climate, more frequent extreme precipitation events have amplified the risk of ERDEs in the YRB (Yin et al., 2018; Yin et al., 2021). Consequently, developing effective early-warning technologies for ERDEs is crucial for safeguarding water security in the region.

Precipitation is the most direct trigger of ERDEs (Wright et al., 2020; Wang et al., 2021; Breinl et al., 2021). Observational analyses show that heavy or persistent rainfall typically occurs within five days before an ERDE (Xu et al., 2018; Tuel et al., 2022; Deng et al., 2022; Xu et al., 2026). Many studies have attempted to identify robust forecasting predictors by examining the meteorological processes that lead to ERDEs (Fang et al., 2018; Cui et al., 2022; Yang et al., 2022). Local weather systems—including the plateau vortex, southwest vortex, and Jianghuai cyclone—are recognized as important indicators for heavy rainfall and

* Corresponding author at: School of Geographical Sciences, Southwest University, Chongqing 400715, China.
E-mail address: liyao7@swu.edu.cn (Y. Li).

https://doi.org/10.1016/j.jag.2026.105590
Received 9 March 2026; Received in revised form 5 September 2026; Accepted 11 September 2026
Available online 22 September 2026
1569-8432/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

heightened ERDE risk (Ding and Li, 2017; Zeng and Wang, 2022; Li and Zhang, 2023; Zhu and Chen, 2023). However, not all precipitation leads to ERDEs. ERDEs occur only when specific atmospheric signals interact with local forcing factors, such as topography, soil moisture, and human interventions (Wang et al., 2022; Kumar et al., 2023). Thus, constructing an early-warning model grounded in clear physical mechanisms remains a key bottleneck in improving ERDE prediction.

Existing ERDE forecasting methods are generally categorized into numerical models and data-driven artificial intelligence models (Breinl et al., 2021). Numerical models simulate the physical processes of ERDE development (Liu et al., 2018), enabling long-term projections under different climate scenarios by imposing initial conditions and external forcing factors (Try et al., 2022; Hodnebrog et al., 2022; Liu et al., 2025). At sub-seasonal to daily scales, numerical models focus on detecting precursory atmospheric circulation signals (Xie et al., 2021; Ding et al., 2024). For example, an eastward-propagating Rossby wave can induce an anomalous low pressure over the eastern Tibetan Plateau, ultimately leading to heavy precipitation in the Sichuan Basin one week later (Nie and Sun, 2021). Achieving accurate sub-daily forecasts, however, requires convective-permitting resolutions (1–4 km) and sophisticated physical parameterizations that capture the effects of complex terrain, leading to high computational costs and substantial uncertainties (Prein et al., 2017; Bližňák et al., 2019; Tang et al., 2023).

In contrast, machine learning models provide significant computational efficiency and have attracted increasing attention for ERDE forecasting (Bi et al., 2023; Kafi and Ponrahono, 2024). Recent breakthroughs in AI-driven forecasting have demonstrated state-of-the-art performance in global weather prediction, probabilistic extreme-weather forecasting, and large-scale flood forecasting. Notable examples include Pangu-Weather, GraphCast, GenCast, and AI-based global flood forecasting systems for ungauged watersheds (Bi et al., 2023; Lam et al., 2023; Nearing et al., 2024; Price et al., 2025). Spatial feature learning models (e.g., convolutional neural networks and attention-based architectures) have been applied to learn the distribution of flood indices, aiming to forecast ERDEs (Davenport and Diffenbaugh, 2021; Adikari et al., 2021; Cui et al., 2025). Temporal machine learning models (e.g., long short-term memory networks, recurrent neural networks, and generative adversarial networks) have been widely used for river discharge and runoff forecasting, providing future information to support ERDE projection (Xie et al., 2022; Zhao et al., 2022). Some studies integrate both spatial and temporal features to build more complex deep learning flood prediction systems (Adikari et al., 2021; Zhang et al., 2022). However, despite these advances, machine learning models still rely heavily on the quantity and quality of training data (Frame et al., 2022). Their performance in forecasting low-probability extreme events such as ERDEs is often hampered by limited event samples, leading to overfitting, compromised precision, and high false-negative rates (Sillmann et al., 2017; de Sousa Araújo et al., 2022). Moreover, purely data-driven models typically lack explicit integration of physical processes, weakening their interpretability and physical consistency (Mukkavilli et al., 2023; Charlton-Perez et al., 2024). Therefore, it is crucial to develop a machine learning framework that incorporates physical mechanisms and maintains robust performance under limited training data.

To address these limitations, this study aims to detect atmospheric circulation precursor signals associated with ERDEs in the YRB. We develop a physically guided DetRF framework that integrates anomaly detection with random forest classification. Specifically, the anomaly detection module extracts abnormal large-scale circulation states, while the random forest classifier determines whether these anomalies are conducive to ERDE occurrence. The detected precursor signals are independently evaluated using ERA5 reanalysis, precipitation, and high-resolution daily river discharge data from 2020 to 2024, thereby establishing a direct connection between atmospheric circulation precursors and subsequent hydrological extremes. The key novelty of this study lies in the development of a physics-guided machine learning framework that identifies physically interpretable, large-scale circulation anomalies linked to ERDE-favorable conditions. Unlike conventional data-driven river discharge forecasting models, DetRF provides a hybrid physics-guided approach for ERDE precursor detection, offering both improved interpretability and enhanced robustness under limited training data.

2. Data and methods

2.1. Datasets

This study used five datasets for analysis and modeling:

(1) ERA5 reanalysis data: Training, validation, and test sets were derived from the fifth-generation ECMWF reanalysis (ERA5) at 0.25° spatial and daily temporal resolution for 2000–2024, with 2020–2024 reserved for an independent application period. Atmospheric inputs included zonal wind, meridional wind, and geopotential heights at 200, 500, 700, and 850 hPa (Hersbach et al., 2023).
(2) Precipitation data: Precipitation was obtained from the half-hourly final run of GPM-IMERG V06 (0.1° × 0.1°) for 2000–2024 (Hou et al., 2014). The data were aggregated to daily cumulative precipitation.
(3) River discharge data: High-resolution daily river discharge data for China (0.05° × 0.05°) from the Global Flood Awareness System (GloFAS) reanalysis for 2020–2024 were obtained from the Joint Research Centre (JRC) of the European Commission (Harrigan et al., 2020). These data were used to independently evaluate the early-warning performance of the DetRF framework for ERDE detection during the application period.
(4) DEM data: Terrain information was derived from the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) database v4.1 at 3 arc-second resolution (~90 m) (Jarvis et al., 2008), accessed via Google Earth Engine.
(5) River-order data: River order was extracted from the MERIT_Hydro_v07_Basins v01 dataset, a global vector hydrographic database derived from the 90 m MERIT-Hydro product (Yamazaki et al., 2019).

2.2. Study area

The YRB spans approximately 1.8 million km^2 across southern China. The terrain descends sharply from west to east, forming an elevation gradient exceeding 5,000 m (Fig. 1a). This vast and complex topography contributes to substantial spatial heterogeneity in ERDEs. Following Xu et al. (2026), the nine subregions were delineated based on transitions in river slope and terrain elevation derived from MERIT Hydro and SRTM. Local precipitation was calculated within each target subregion, whereas upstream precipitation was calculated within the immediately adjacent, hydrologically connected upstream subregion.

Fig. 1b shows the cumulative number of daily ERDE occurrences during 2000–2019. The Dongting Lake Plain (Reg7) recorded the highest number of ERDEs (482), followed by the Central-eastern Sichuan Basin (Reg5) with 365 events. Most subregions—such as the Western Sichuan Plateau (Reg3), Western Sichuan Basin (Reg4), Central-eastern Sichuan Basin (Reg5), Wushan Mountain Range (Reg6), Dongting Lake Plain (Reg7), and Poyang Lake Plain (Reg8)—experienced more than 200 ERDEs. In contrast, fewer than 100 ERDEs occurred in the YRB source region (Reg1), eastern Tibetan Plateau (Reg2), and Lower Yangtze River Plain (Reg9). However, the five-day cumulative precipitation preceding ERDEs was highest in the Sichuan Basin and Daba Mountains (Reg4–Reg6), moderate in the mid-lower reaches (Reg8–Reg9), and lowest in the upper reaches (Reg1–Reg3). This discrepancy indicates a nonlinear relationship between ERDE frequency and antecedent precipitation.

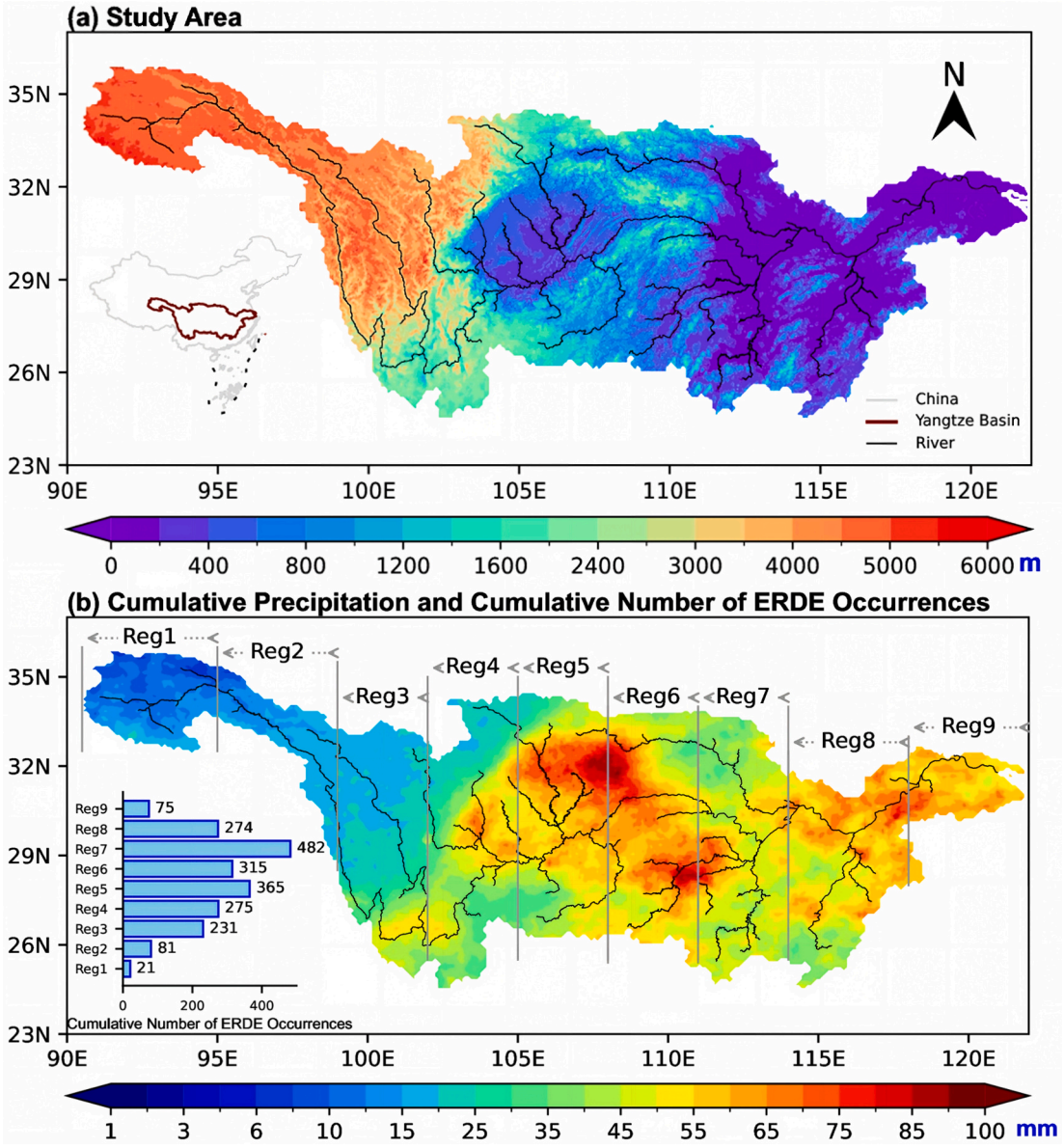


Fig. 1. (a) Terrain elevation (shading, unit: m) and spatial distribution of rivers with a river order of four or higher based on the MERIT dataset. (b) Spatial distribution of five-day cumulative precipitation preceding ERDEs (shading, unit: mm) and the cumulative number of daily ERDEs during 2000–2019 in each subregion. Grey lines denote the boundaries of the nine subregions (Reg1–Reg9).

Based on the findings of Xu et al. (2026), ERDE-related precipitation patterns in the YRB can be classified into three categories: (1) upstream-type ERDEs, primarily driven by upstream precipitation (Reg2 and Reg7); (2) local-type ERDEs, mainly controlled by local precipitation (Reg4 and Reg8); and (3) overlapped-type ERDEs, resulting from the combined effects of upstream and local precipitation (Reg3, Reg5, Reg6, and Reg9). Reg1 was excluded from the analysis because it lacks an upstream contributing area. Xu et al. (2026) further investigated the meteorological mechanisms underlying these three ERDE types and identified the characteristic atmospheric circulation signals associated with each category. Building upon these insights, we developed the DetRF framework, a machine learning model designed to automatically identify predefined atmospheric circulation precursor signals, thereby improving the early warning of ERDEs in the YRB.

2.3. Methods

(1) Data Preparation

The model inputs were derived from three key meteorological variables that represent the predefined atmospheric circulation signals associated with ERDEs: geopotential height, zonal wind, and meridional wind. Based on the pressure levels at which the atmospheric circulation signals occur (200, 500, 700, and 850 hPa), meteorological variables were extracted from the ERA5 reanalysis dataset for the period March–November during 2000–2024. Each variable was independently normalized using its corresponding minimum and maximum values, and the normalized variables across all pressure levels were then combined to construct the multidimensional feature matrix **X**. Binary labels were generated based on the event-to-precursor temporal mapping following Xu et al. (2026). Let T_e denote the ERDE occurrence time and T_s represent the representative time of its associated atmospheric circulation precursor. The precursor time was defined as $T_s = T_e - L_r$, with a lead time $L_r = 120h$ for Reg2–Reg5 and Reg7–Reg9, and $L_r = 72h$ for Reg6. These regional lead times correspond to periods when the relevant weather systems were already established prior to ERDE occurrence. The 3-hourly precursor timestamps were subsequently aggregated into

daily labels. A day was labeled as positive (Y = 1) if it contained at least one identified precursor timestamp T_s, and as negative (Y = 0) otherwise. To avoid overrepresenting persistent synoptic systems, only one reference precursor sample was retained for each ERDE. Consequently, the number of positive circulation samples equaled the number of ERDEs, although their dates preceded the corresponding ERDEs by the prescribed regional lead times.

The feature matrix **X** and label **Y** for the period March–November during 2000–2019 were used for model training and evaluation. The feature variable **X** from 2020 to 2024 was reserved as an independent application dataset to assess the model's ability to detect atmospheric circulation signals relative to the predefined circulation patterns. The predictive skill of the detected signals was then independently evaluated using precipitation and reanalysis river discharge data, thereby assessing the framework's early-warning capability for ERDEs (Fig. 2a).

(2) Model construction

As shown in Table 1, positive samples accounted for less than 10% of the data in each subregion, indicating a highly imbalanced classification problem under data-scarce conditions. To address this challenge, DetRF integrates an isolation forest for anomaly detection with a balanced random forest for classification. Eight independent DetRF models were constructed for subregions Reg2–Reg9 (Fig. 2b).

The isolation forest is an efficient unsupervised anomaly detection algorithm well suited for high-dimensional data (Liu et al., 2008). It identifies anomalies based on how easily a sample can be isolated from the rest of the dataset (Rawal and Patel, 2025). In this study, predefined atmospheric circulation signals were treated as anomalous instances (signal days). The algorithm was trained on common circulation fields (non-signal days) to learn their typical characteristics, and the resulting anomaly scores were incorporated as an additional feature in the training dataset to enrich the feature representations for subsequent classification. The isolation forest model was configured with 200 trees (n_estimators = 200), a value determined through preliminary sensitivity analyses. The contamination parameter was specified separately for each subregion based on the proportion of positive samples in the training dataset, thereby preventing information leakage from the validation periods.

The feature-augmented data were then input into a balanced random forest classifier, which mitigates class imbalance through balanced bootstrap sampling, ensuring equal representation of both classes in each decision tree's training sample (Xie et al., 2009). The number of trees was also set to 200. Through parameter tuning, the sampling strategy was adjusted such that the proportion of positive samples in each bootstrap sample was approximately 0.5.

(3) Model evaluation

The DetRF model was first evaluated using K-fold cross-validation, in which the dataset was randomly partitioned into five subsets (Kohavi, 1995). In each iteration, four subsets were used for training, with the remaining subset serving as the validation set, allowing each subset to be used once for validation. This procedure provides a robust and unbiased estimate of DetRF's performance. Because daily atmospheric fields may exhibit temporal dependence, we conducted an additional year-wise temporal validation experiment. For each subregion, the most recent year containing positive signal samples during 2000–2019 was selected as the validation year, while all preceding years were used for model training. Feature normalization, isolation forest anomaly score estimation, and balanced random forest training were performed exclusively on the training years. No samples from the validation year were involved

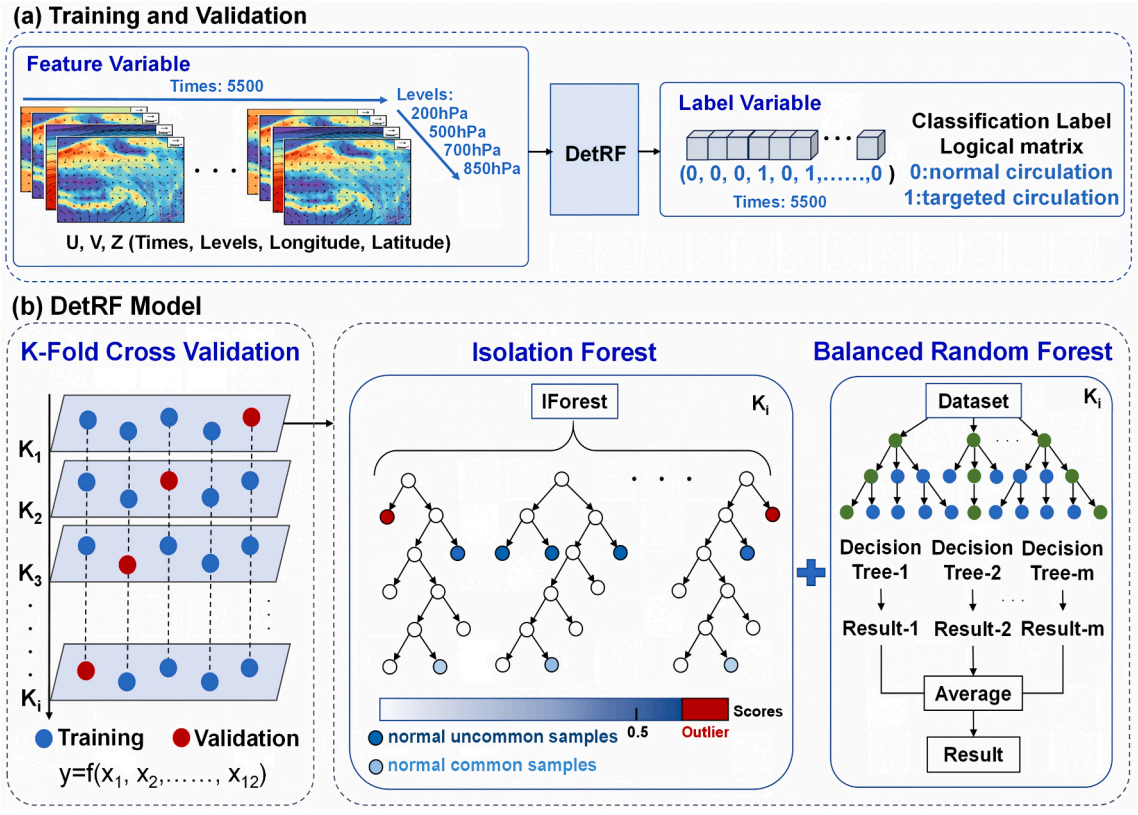


Fig. 2. Schematic diagram of the DetRF model architecture. (a) Illustration of the input feature matrix **X** and the binary label variable **Y** during the training and validation period. (b) The DetRF architecture consists of a sequential framework that integrates a K-fold cross-validation training strategy, an isolation forest algorithm for identifying anomalous samples in the feature space, and a balanced random forest algorithm that mitigates class imbalance by training each decision tree on a balanced bootstrap sample.

Table 1
Distribution of negative samples, positive samples, and positive sample ratios across subregions (Reg2–Reg9) during 2000–2019.

	Reg2	Reg3	Reg4	Reg5	Reg6	Reg7	Reg8	Reg9
Negative	5419	5269	5225	5135	5185	5018	5226	5425
Positive	81	231	275	365	315	482	274	75
Positive Ratio	1.473%	4.200%	5.000%	6.636%	5.727%	8.764%	4.982%	1.364%

in preprocessing, model fitting, hyperparameter selection, or threshold determination. This design effectively prevents leakage of temporal information from the validation year and provides a more conservative and realistic assessment of model performance under temporally independent validation conditions.

Model performance was evaluated using precision, recall, F1-score (Provost and Fawcett, 2001), and the False Negative Rate (FNR) (Hastie et al., 2009), defined as follows:

$$precision = \frac{TP}{TP + FP} \tag{1}$$

$$recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \tag{3}$$

$$FNR = \frac{FN}{TP + FN} \tag{4}$$

These metrics were computed from the confusion matrix using true positives (TP), false positives (FP), and false negatives (FN). Precision represents the proportion of predicted positive samples that are truly positive, while recall indicates the proportion of actual positive samples correctly identified. The F1-score provides the harmonic mean of precision and recall. Given the critical consequences of missing rare ERDEs, the FNR is included to quantify the proportion of undetected positive events, thereby reflecting the model's omission error.

The DetRF model outputs the probability of occurrence of a predefined atmospheric circulation signal at each time step. To determine the optimal classification threshold, precision-recall (PR) curves were constructed for each subregion (Fig. A1). Precision, recall, and F1-score were calculated across a range of probability thresholds. Because precision is particularly sensitive to false positives, PR curves provide an effective framework for evaluating the trade-off between signal detection and false-alarm control. Based on this analysis, a probability threshold of 0.86 was selected as a relatively conservative criterion that substantially reduces false positives while maintaining satisfactory recall (Fig. A2). To ensure consistency and comparability across regions, the same threshold (0.86) was applied uniformly to all subregions.

3. Results and discussion

3.1. Quantitative evaluation of the DetRF model

The validation results of the DetRF model, obtained via random K-fold cross-validation for 2000–2019, are summarized in Table 2. Overall, DetRF demonstrated robust and consistent performance in detecting atmospheric circulation signals associated with ERDEs across all YRB subregions. The model exhibited strong detection capability, with recall values exceeding 0.815 across all regions, peaking at 0.962 in Reg7. These results indicate that DetRF successfully identified the majority of predefined atmospheric circulation signals, suggesting a low rate of missed detections. Correspondingly, the FNR remained consistently low across the basin, ranging from 0.037 (Reg7) to 0.185 (Reg2), confirming the model's effectiveness in minimizing missed positive events. Precision values exhibited minimal regional variation, clustering within a narrow range from 0.502 (Reg5) to 0.509 (Reg3), suggesting that the model maintained stable positive predictive performance despite spatial heterogeneity. The F1-score, which balances precision and recall, ranged from 0.626 (Reg2) to 0.661 (Reg3), indicating a well-balanced overall performance. These validation results demonstrate the strong performance and reliability of DetRF in detecting ERDE-related atmospheric circulation signals, supporting its application to the independent period for further assessment of operational early-warning performance.

Table 2
Performance evaluation of the DetRF model based on random K-fold cross-validation across YRB subregions during March–November of 2000–2019.

Region	Recall	FNR	Precision	F1-score
Reg2	0.815	0.185	0.508	0.626
Reg3	0.939	0.061	0.509	0.661
Reg4	0.927	0.073	0.505	0.654
Reg5	0.943	0.056	0.502	0.655
Reg6	0.946	0.054	0.503	0.656
Reg7	0.962	0.037	0.503	0.661
Reg8	0.945	0.055	0.504	0.657
Reg9	0.867	0.133	0.508	0.640

In addition, precision was approximately 0.500 across all YRB subregions and should therefore be interpreted with caution. The DetRF framework was designed to identify atmospheric signals associated with ERDEs and to function as a preliminary signal-screening tool, rather than to provide a deterministic diagnosis of ERDE occurrence based solely on the prediction-day label. To further investigate the nature of false-positive detections, a lead-event analysis was conducted using only the validation folds. For each false-positive sample, we examined whether a positive label occurred within the subsequent five days and compared this proportion with that of true-negative samples. The results showed that a fraction of false-positive samples were followed by positive labels within five days, particularly in Reg3, Reg6, and Reg8, with corresponding proportions of 15.38%, 15.38%, and 11.76%, respectively. In contrast, the proportions for true-negative samples in the same regions were substantially lower, at 1.07%, 1.92%, and 1.42%, respectively.

These findings suggest that some samples classified as false positives may represent atmospheric conditions that subsequently evolved into ERDE-favorable circulation states. However, this effect was not consistently observed across all regions, and the overall proportion remained relatively small. Consequently, the evidence does not support reclassifying false positives as successful detections. Instead, false positives should be interpreted conservatively as a mixture of genuine false alarms and potential early-stage atmospheric signals that precede the occurrence of predefined circulation patterns. Despite this limitation, DetRF achieved precision values that exceeded those reported for machine learning applications developed under similarly data-scarce conditions (Wang et al., 2024; Camps-Valls et al., 2025).

Under the year-wise temporal validation, the basin-wide mean precision and F1-score for the final validation year were 0.506 and 0.576, respectively. The reduction in model performance exhibited substantial spatial heterogeneity. Reg2, Reg3, and Reg9 maintained relatively high predictive skill, whereas Reg4–Reg8 showed more pronounced declines in precision and F1-score. These differences are physically plausible because ERDEs in Reg4–Reg8 are influenced by a wider range of synoptic systems and local hydrometeorological processes, including forward-tilting troughs, southwest vortices, double vortices, backward-tilting troughs, and Jianghuai cyclones. The diversity of these circulation systems in terms of spatial scale, vertical structure, and persistence

increases the complexity of distinguishing signal from non-signal conditions, thereby making the binary classification task more challenging.

Despite the reduction in performance relative to random K-fold cross-validation, the year-wise temporal validation results indicate that DetRF retains useful capability for detecting ERDE-related atmospheric circulation signals under a more stringent and temporally independent evaluation framework. Given that the primary objective of DetRF is to serve as a precursor-screening tool rather than a deterministic forecasting model, the performance obtained from random K-fold cross-validation was considered acceptable for subsequent analyses. Therefore, we retained the DetRF configuration with random K-fold cross-validation as the operational atmospheric circulation signal detection model for ERDE precursor screening (Table 3).

3.2. Detection of atmospheric circulation signals using the DetRF model

The performance of the DetRF model in detecting predefined atmospheric circulation signals associated with ERDEs was evaluated using independent daily ERA5 data from March to November during 2020–2024. The detected signals were compared with the predefined signals from 2000 to 2019 across the eight subregions (Reg2–Reg9).

For upstream-type ERDEs in Reg2, the predefined atmospheric circulation signals featured a distinct 500-hPa plateau vortex centered near 34–36°N and 88–90°E in the western upstream region. Dynamically, this upstream plateau vortex interacted with the South Asian High at 200 hPa, generating strong upper-level divergence and enhancing low-level air ascent (Fig. 3a). The DetRF model accurately captured the major characteristics of this signal during 2020–2024. The model-detected signals exhibited a stronger plateau vortex, a northward-extended South Asian High, and intensified low-level southwesterly convergence over Reg2 and its upstream region compared with the predefined benchmark (Fig. 3b). Despite the limited number of event samples, the detected signals exhibited a clearer atmospheric circulation signature.

A similar result was found for Reg7, where upstream precipitation was the dominant driver of ERDEs. The predefined atmospheric circulation signal consisted of a backward-tilting trough at 700-hPa near 104°E in the western upstream area of Reg7, sloping southwestward with height. Strong 700-hPa southwesterlies ahead of the trough transported moisture into Reg7 and its upstream regions (Fig. 3c). The DetRF model accurately detected this synoptic configuration during 2020–2024, capturing both the backward-tilting trough within 24–34°N and 102–108°E and the associated moisture transport (Fig. 3d). These results confirm that DetRF effectively identifies the critical upstream atmospheric circulation structures responsible for ERDEs, underscoring its value for early warning applications.

For local-type ERDEs in Reg4, the primary atmospheric circulation signal was a forward-tilting trough sloping northeastward with height. Southerly winds east of the trough transported substantial moisture into Reg4, where the influence of the Tibetan Plateau enhanced moisture convergence and uplift, triggering localized precipitation (Fig. 4a). DetRF successfully captured the spatial structure of this trough within 25–35°N and 96–104°E over western Reg4, as well as the associated 700-hPa southerly wind with moisture transport during the independent application period (Fig. 4b). A similar pattern was observed for Reg8, where local precipitation was the dominant ERDE trigger. DetRF effectively detected the Jianghuai cyclone within Reg8 (29–33°N, 112–116°E) during 2020–2024 (Fig. 4c). The model-detected cyclone exhibited slightly weaker intensity and southwesterly winds relative to the predefined atmospheric circulation signal (Fig. 4d), likely reflecting the limited number of events available during the independent application period. Nevertheless, the model accurately reproduced the dominant southwesterly moisture transport pathway, which is critical for ERDE development. These results demonstrate the robust performance of DetRF in detecting synoptic weather systems associated with ERDEs in Reg4 and Reg8.

Table 3
Performance evaluation of the DetRF model based on year-wise temporal validation across YRB subregions during March–November of 2000–2019.

Region	Recall	FNR	Precision	F1-score
Reg2	0.852	0.148	0.622	0.719
Reg3	0.892	0.108	0.831	0.860
Reg4	0.500	0.500	0.357	0.417
Reg5	0.808	0.192	0.333	0.472
Reg6	0.968	0.032	0.297	0.455
Reg7	0.500	0.500	0.579	0.537
Reg8	0.755	0.245	0.298	0.428
Reg9	0.769	0.231	0.727	0.748

For overlapped-type ERDEs in Reg3, the predefined atmospheric circulation signals featured a 500-hPa westerly trough (34–36°N, 96–102°E) coupled with a South Asia High at 200 hPa. This synoptic configuration enhanced upper-level divergence, promoted convergence east of the trough, and strengthened southwesterly winds that transported moisture into Reg3 and the Reg2–Reg3 boundary (Fig. 5a). The DetRF model successfully captured this pattern during 2020–2024, detecting a deepened westerly trough and a more expansive South Asian High compared with the predefined benchmark (Fig. 5b).

In Reg5, the predefined atmospheric circulation signals exhibited a distinct 850-hPa southwest vortex near the Reg4–Reg5 boundary, accompanied by the Western North Pacific Subtropical High (WNPSH) extending across 22–30°N and 102–114°E, which facilitated southerly moisture transport into the Reg4–Reg5 boundary region (Fig. 5c). The DetRF model accurately reproduced this structure, detecting the southwest vortex at the Reg4–Reg5 junction and capturing a stronger vortex and a more expansive WNPSH relative to the predefined signals (Fig. 5d).

Similarly, the predefined atmospheric circulation signals in Reg6 consisted of a double-vortex system within 24–27°N and 101–104°E, along with the WNPSH over southeastern Reg6, together supporting persistent southerly moisture transport (Fig. 6a). The DetRF model successfully detected the double-vortex pattern during the independent application period. A slight northeastward shift in one vortex was observed, and the model produced a stronger representation of the WNPSH than the predefined benchmark (Fig. 6b). These differences suggest that, while DetRF effectively identified the dominant synoptic features shaped by the Sichuan Basin topography, finer-scale spatial details showed minor deviations. In Reg9, overlapped-type ERDEs were associated with a northward extension of the WNPSH to 32°N, producing prevailing southwesterly winds over Reg9 and its upstream region (Fig. 6c). The DetRF model accurately captured these atmospheric circulation signals, including the intensified WNPSH ridge and the associated low-level moisture transport (Fig. 6d).

Overall, the DetRF model reliably captured the key spatial characteristics of predefined atmospheric circulation signals, including the locations of troughs, vortices, and cyclones, as well as the associated moisture transport. Although slight discrepancies in intensity or spatial positioning were observed for some systems, these were likely due to the limited number of independent samples available for 2020–2024. Across all subregions, the spatial correlations between the predefined atmospheric circulation signals and the composite DetRF-detected signals ranged from 0.702 to 0.995, with a mean of 0.915, and all correlations were statistically significant at the 95% confidence level ($p < 0.05$) (Table A1). Nevertheless, the strong agreement between predefined and detected signals demonstrates that DetRF is a robust and effective tool for identifying atmospheric circulation signatures associated with ERDEs.

3.3. Assessing the predictability of ERDEs from precipitation

To assess the predictability of ERDEs, we first analyzed precipitation events within five days following the detection of atmospheric circulation signals by the DetRF model. The five-day evaluation window was

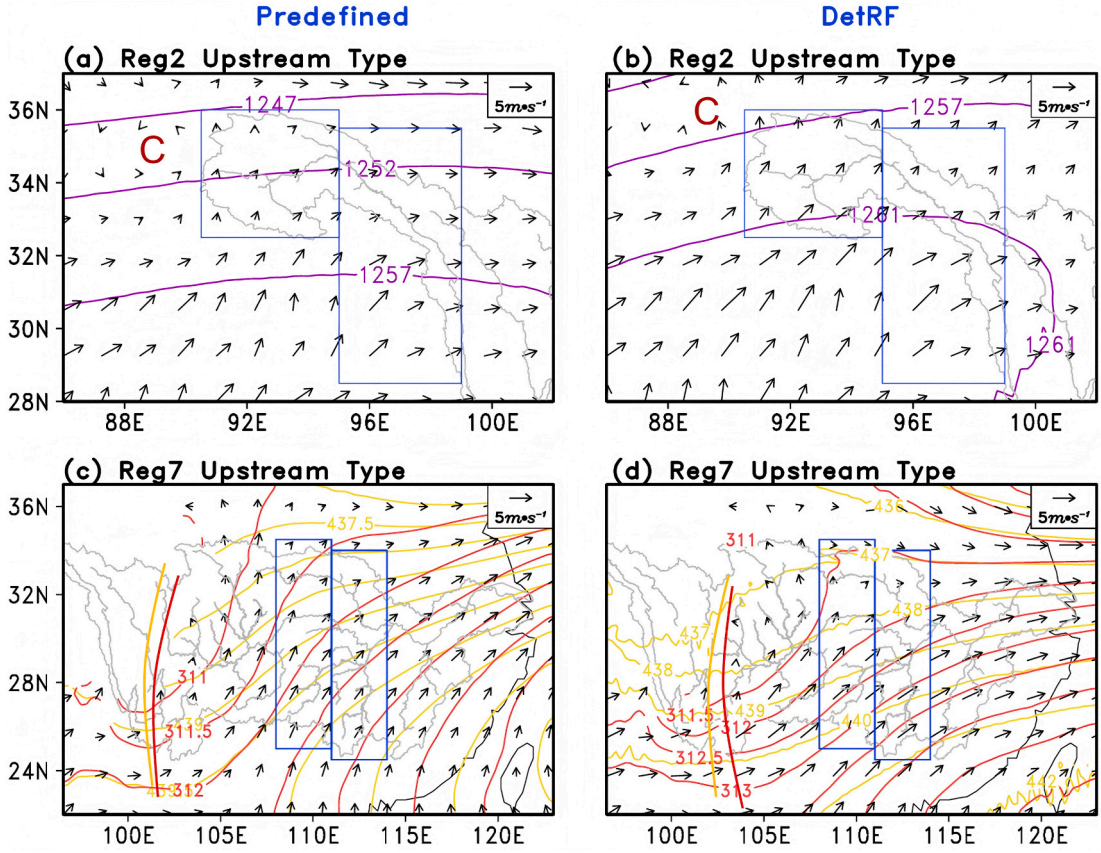


Fig. 3. Comparison between predefined atmospheric circulation signals (2000–2019) and DetRF-detected signals (2020–2024) for the upstream-type ERDE regions, Reg2 and Reg7. (a–b) Spatial distributions of 200-hPa geopotential height (purple contours; unit: dagpm) and 500-hPa wind fields (vectors; unit: m/s) over Reg2. The red "C" denotes the plateau vortex. (c–d) Spatial distributions of 600-hPa (yellow contours) and 700-hPa (red contours) geopotential heights, along with 700-hPa wind fields (vectors; unit: m/s) over Reg7. The yellow and red solid lines represent the backward-tilting trough at 600 and 700 hPa, respectively. The blue box outlines the target subregion and its upstream area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

defined as the period from the signal day (day 0) through day +5 after signal detection, rather than as a fixed five-day lead time. We examined the occurrence of precipitation events (daily mean precipitation > 0.1 mm), persistent precipitation events (≥2 consecutive days with precipitation), and extreme precipitation days (daily cumulative precipitation exceeding the 90th percentile of the historical records from March to November 2020–2024 over each subregion).

For upstream-type ERDEs, DetRF demonstrated strong predictive skill in precipitation by detecting the atmospheric circulation signals. In Reg2, all detected signals were followed by precipitation and persistent precipitation events, while extreme precipitation occurred on 38.6% and 55.0% of rainy days in the local and upstream areas, respectively (Fig. 7a). In Reg7, more than 99% of detected signals were followed by precipitation, with over 89% resulting in persistent precipitation. Extreme precipitation accounted for 11.6% and 12.1% of rainy days in Reg7 and its upstream area, respectively (Fig. 7f).

For local-type ERDEs, DetRF also showed high predictive performance. In Reg4, all detected signals were followed by precipitation, with >96% resulting in persistent precipitation, while extreme precipitation comprised 21.2% of rainy days (Fig. 7c). In Reg8, all detected signals were followed by both precipitation and persistent precipitation, with extreme precipitation accounting for 23.5% of rainy days (Fig. 7g).

For overlapped-type ERDEs, the model exhibited generally strong but regionally variable predictive capability. In Reg3, Reg5, and Reg6, all detected signals were followed by precipitation, with persistent precipitation occurring in more than 86% of events in both local and upstream areas. Extreme precipitation accounted for 16.9–28.6% of rainy days (Figs. 7b, d, and e). In contrast, Reg9 exhibited distinct characteristics, with lower frequencies of persistent precipitation (77.8% locally and 66.7% upstream) and substantially lower proportions of extreme precipitation days (2.1% locally and 3.1% upstream) (Fig. 7h).

Overall, the DetRF model effectively captured the occurrence, persistence, and intensity of precipitation within five days following the detection of atmospheric circulation signals across the YRB during 2020–2024. These findings demonstrate DetRF's capability to provide reliable early warnings of ERDEs by detecting atmospheric circulation precursors.

Fig. 8 illustrates the spatial distribution of cumulative precipitation within five days following the detection of atmospheric circulation signals across YRB subregions during 2020–2024. For upstream-type ERDEs in Reg2 and Reg7, the cumulative precipitation maxima were predominantly concentrated in their respective upstream areas (Reg1 and Reg6), exhibiting a clear downstream transport pattern along the river channel (Figs. 8a and f). This spatial pattern aligned with the upstream atmospheric circulation signals detected by the DetRF model. For local-type ERDEs in Reg4 and Reg8, cumulative precipitation was primarily confined within the target subregions, with pronounced centers in central Reg4 and southern Reg8 (Figs. 8c and g). These distributions highlight the effectiveness of the detected atmospheric circulation signals in predicting the local precipitation patterns. For overlapped-type ERDEs in Reg5 and Reg6, multiple precipitation centers developed across both upstream and local areas, reflecting the influence of weather systems positioned along boundaries between adjacent subregions (Figs. 8d and e). In Reg3, cumulative precipitation was concentrated locally, likely due to the rapid eastward progression of the associated

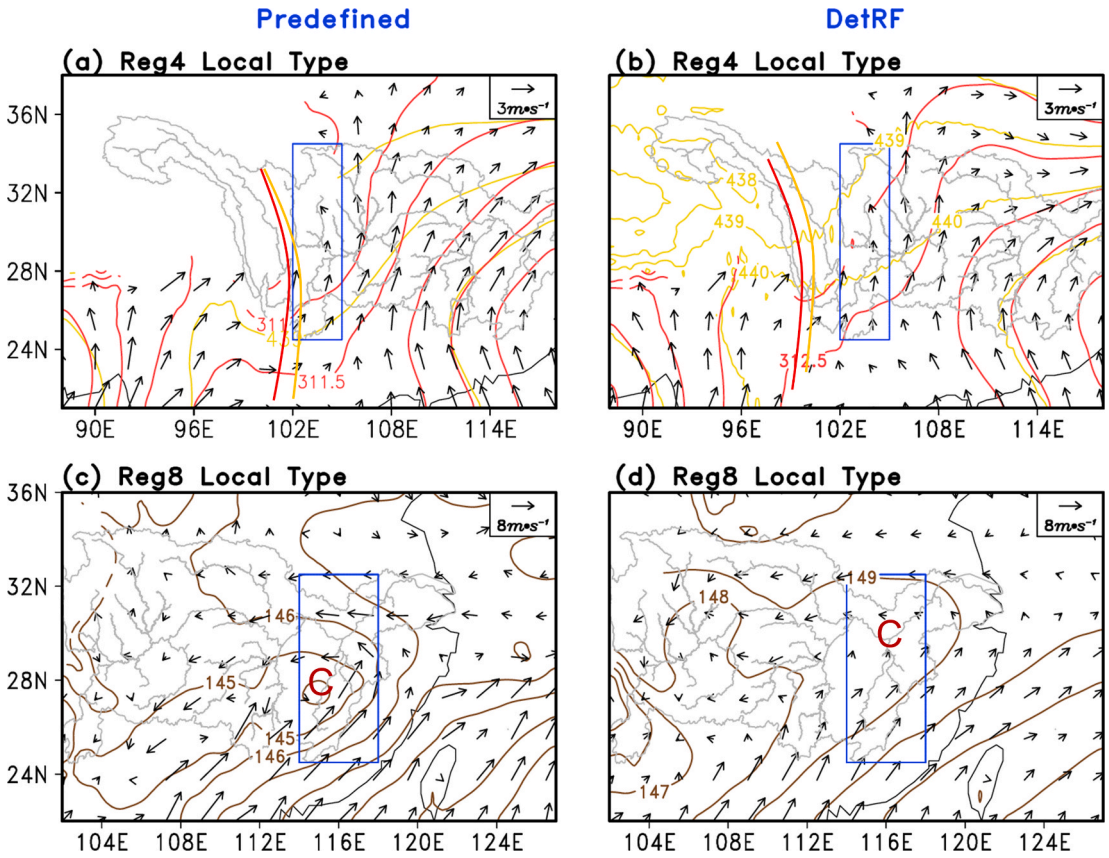


Fig. 4. Comparison between predefined atmospheric circulation signals (2000–2019) and DetRF-detected signals (2020–2024) for the local-type ERDE regions, Reg4 and Reg8. (a–b) Spatial distributions of 600-hPa (yellow contours; unit: dagpm) and 700-hPa (red contours) geopotential heights, along with 700-hPa wind fields (vectors; unit: m/s) over Reg7. The yellow and red solid lines denote the forward-tilting troughs at 600 and 700 hPa, respectively. The blue box outlines the target subregion. (c–d) Spatial distributions of 850-hPa geopotential height and wind fields over Reg8. The red "C" indicates the Jianghuai cyclone. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

westerly trough (Fig. 8b). In Reg9, convective precipitation linked to the western flank of the WNPSH exhibited a relatively small spatial footprint, short duration, and high intensity (Fig. 8h). Although the affected area was smaller compared with other regions, distinct precipitation centers were still observed in eastern Reg9 and the central-southeastern portion of its upstream area. Overall, the atmospheric circulation signals detected by the DetRF model effectively captured the spatial patterns of precipitation associated with ERDEs, highlighting their potential for early warning applications.

3.4. *Evaluating the predictability of ERDEs using river discharge data*

We further evaluated the predictability of ERDEs using high-resolution daily river discharge reanalysis data during the independent application period. To ensure consistency with the ERDE definition, an ERDE was defined at the grid-cell scale as a daily river discharge exceeding the 95th percentile of discharge for that grid cell during March–November of 2020–2024. For each DetRF-detected atmospheric circulation signal, we examined whether at least one valid grid cell exceeded its discharge threshold within five days of its detection.

Fig. 9 shows the number and proportion of ERDE occurrences within five days after DetRF-detected atmospheric circulation signals. The proportions of ERDE occurrences were 100.0% in Reg2, 100.0% in Reg3, 99.2% in Reg4, 93.7% in Reg5, 91.6% in Reg6, 79.8% in Reg7, 100.0% in Reg8, and 86.1% in Reg9. These results indicate that the vast majority of detected atmospheric circulation signals were followed by ERDE occurrence within the subsequent five days. For the upstream-type regions, all 14 detected signals in Reg2 were followed by ERDEs within five days, and the daily proportion of ERDE occurrences remained consistently high from day 0 to day +5 (Fig. 9a). Reg7 also exhibited clear predictability, with 217 of 272 detected signals followed by ERDEs within five days. The proportion of ERDE occurrences was highest during days 0–2 and gradually decreased thereafter, suggesting a relatively rapid yet spatially distributed hydrological response to the detected atmospheric circulation signals (Fig. 9f).

For the local-type regions, 130 of 131 detected atmospheric circulation signals in Reg4 were followed by ERDEs within five days. The daily proportion of ERDE occurrences remained consistently high, ranging from 97.7% to 99.2% throughout the five-day evaluation window (Fig. 9c). In Reg8, all nine detected signals were followed by ERDEs within five days, although the sample size was limited. The proportion of ERDE occurrences was highest on day 0 and day +1, suggesting a relatively rapid hydrological response to the detected atmospheric circulation signals in this region (Fig. 9g).

The overlapped-type regions (Reg3, Reg5, Reg6, and Reg9) also exhibited substantial predictability. In Reg3, all 122 detected signals were followed by ERDEs within five days, and the daily proportion of ERDE occurrences remained 100% across all lag days, indicating a highly stable linkage between the detected atmospheric circulation signals and subsequent ERDE occurrence (Fig. 9b). In Reg5 and Reg6, 93.7% and 91.6% of detected signals, respectively, were followed by ERDEs within five days, with relatively stable daily occurrence proportions throughout the evaluation window (Figs. 9d and e). Reg9 exhibited a weaker but still evident hydrological response, with 31 of 36 detected signals followed by ERDEs within five days and daily ERDE occurrence proportions ranging from 69.4% to 75.0% (Fig. 9h). We further quantified the number of missed ERDEs during the 2020–2024 independent application period. DetRF detected 840 of the 1083 unique

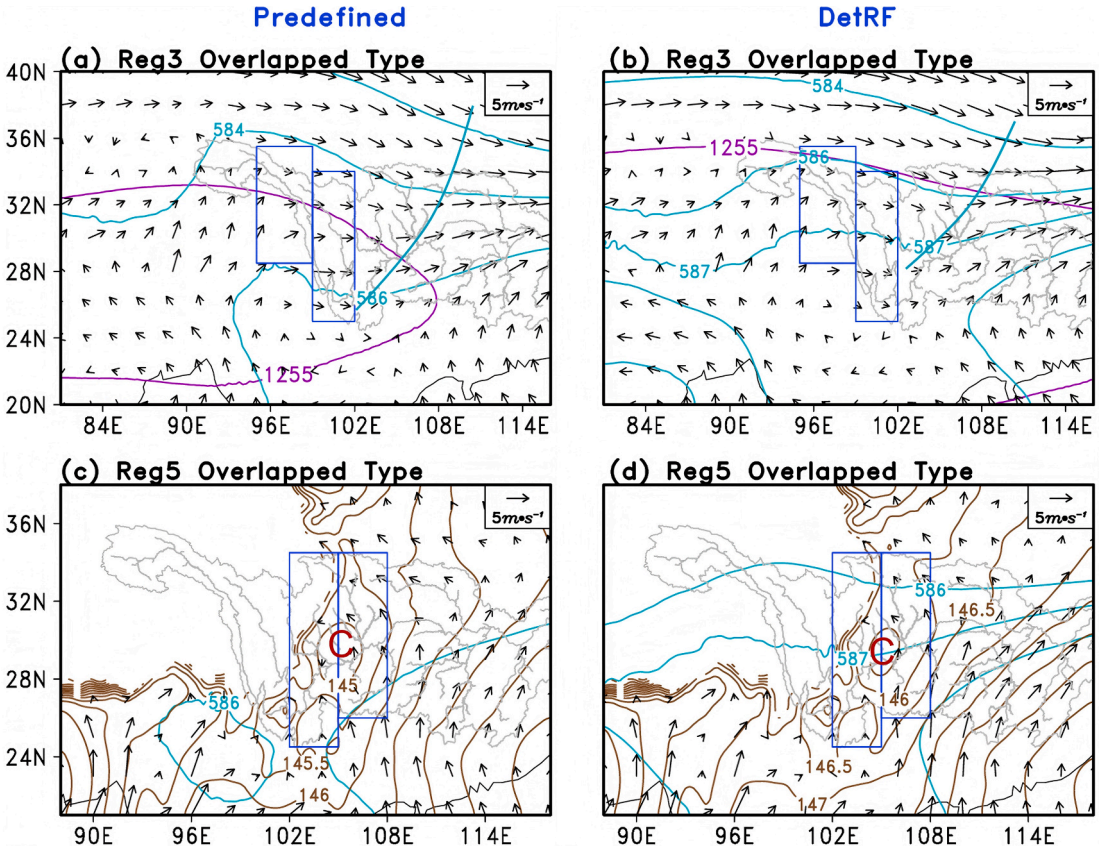


Fig. 5. Comparison between predefined atmospheric circulation signals (2000–2019) and DetRF-detected signals (2020–2024) for the overlapped-type ERDE regions, Reg3 and Reg5. (a–b) Spatial distributions of 200-hPa (purple contours; unit: dagpm) and 500-hPa (light blue contours) geopotential heights, along with 500-hPa wind fields (vectors; unit: m/s) over Reg3. Light blue solid lines denote the westerly trough at 500 hPa. The blue box outlines the target subregion and its upstream area. (c–d) Spatial distributions of 850-hPa (brown contours) and 500-hPa (light blue contours) geopotential heights and wind fields over Reg5. The red "C" indicates the southwest vortex. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ERDE occurrences within the subsequent five-day window, while 243 ERDEs were missed, yielding an overall recall of 77.56% and an FNR of 22.44%. Regional recall ranged from 37.84% in Reg2 to 88.93% in Reg7, with detailed results provided in Table A2.

Fig. 10 further illustrates the spatial distribution of the ERDE occurrence proportion within five days after DetRF-detected atmospheric circulation signals. The basin-wide mean ERDE occurrence proportion was 25.6%, indicating that the detected circulation signals were followed by spatially organized river discharge responses. For the upstream-type regions, Reg2 exhibited the highest proportion of ERDE occurrences, reaching 55.9%, indicating a strong and concentrated hydrological response following signal detection (Fig. 10a). This pattern was consistent with Fig. 8a, where cumulative precipitation was concentrated in the upstream region. In contrast, Reg7 showed a lower maximum ERDE occurrence proportion of 22.2%, suggesting a more spatially dispersed hydrological response (Fig. 10f). This finding agrees with the broader precipitation pattern in Fig. 8f, where precipitation was spread across the upstream area rather than concentrated around a single discharge-response center.

For the local-type regions, Reg4 reached a maximum ERDE occurrence proportion of 35.8%, with high-proportion areas closely corresponding to the local precipitation centers identified in Fig. 8c. This result indicates that atmospheric circulation signals detected by DetRF were followed by both local precipitation accumulation and localized ERDE occurrence (Fig. 10c). Reg8 exhibited an even higher maximum ERDE occurrence proportion of 46.7%, consistent with the extensive precipitation area shown in Fig. 8g and supporting a strong local hydrological response (Fig. 10g).

For the overlapped-type regions, Reg3, Reg5, and Reg6 exhibited relatively high maximum ERDE occurrence proportions of 33.9%, 33.1%, and 37.1%, respectively (Figs. 10b, d, and e). These regions corresponded to the precipitation patterns shown in Fig. 8b, d, and e, where precipitation centers extended across both local and upstream areas, reflecting the combined influence of local rainfall and upstream hydrological contributions. Reg9 showed a somewhat lower but still substantial maximum ERDE occurrence proportion of 30.2%, primarily concentrated along the river network (Fig. 10h). This pattern suggests that localized convective rainfall, mesoscale moisture convergence, dense river–lake networks, and human regulation may influence river discharge responses in Reg9. Such processes were only partially represented by the large-scale atmospheric circulation signals used in DetRF. Future work should incorporate additional predictors, including atmospheric moisture, convective instability, high-resolution precipitation, soil moisture, and hydrological regulation information, to further improve ERDE early warning in Reg9.

Overall, these results demonstrate the strong predictive value of DetRF-detected atmospheric circulation signals as precursors of ERDEs across the YRB. On average, 93.8% of detected signals were followed by at least one ERDE within five days. The resulting ERDE patterns were spatially coherent and physically consistent with the associated precipitation patterns, with a basin-mean ERDE occurrence proportion of 25.6%. The combined temporal and spatial evidence indicates that DetRF effectively identified atmospheric circulation precursors that precede ERDE occurrence over the YRB.

3.5. Comparison of DetRF and benchmark machine learning models

To further evaluate the robustness of DetRF, we conducted a

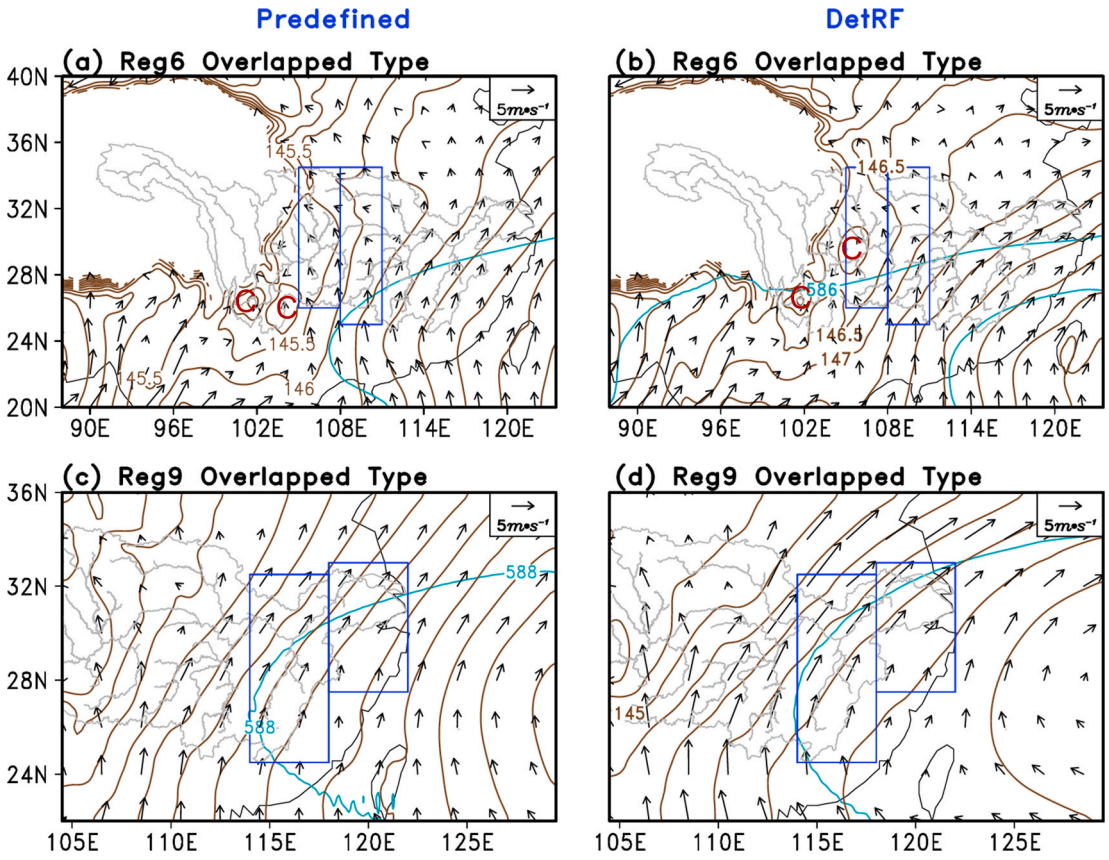


Fig. 6. Comparison between predefined atmospheric circulation signals (2000–2019) and DetRF-detected signals (2020–2024) for the overlapped-type ERDE regions, Reg6 and Reg9. (a–d) Spatial distributions of geopotential height at 500 hPa (light blue contours; unit: dagpm) and 850 hPa (brown contours), along with the 850-hPa wind fields (vector; unit: m/s). The two red "C" symbols indicate the double vortex. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

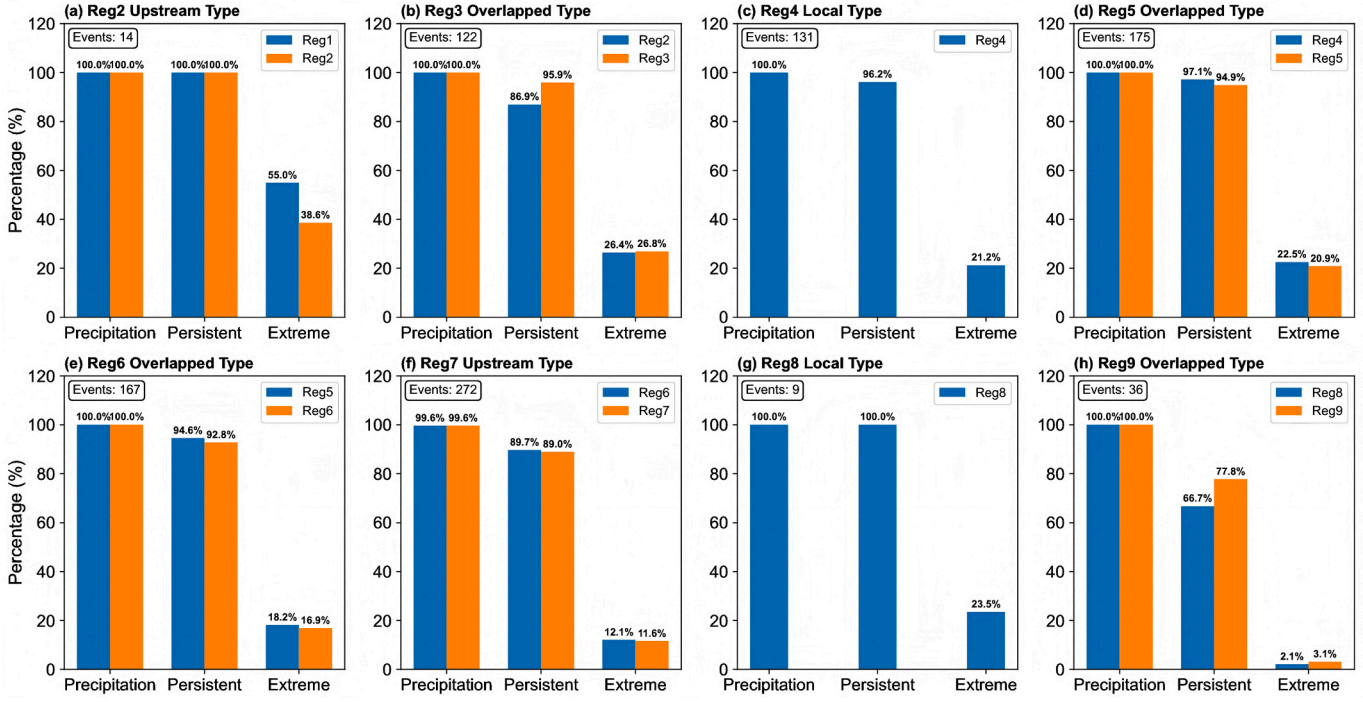


Fig. 7. Histograms showing the proportions of precipitation events (daily mean precipitation > 0.1 mm) and persistent precipitation events (≥2 consecutive days with precipitation) relative to the total number of detected events, and the proportion of extreme precipitation days relative to the total number of rainy days within five days following signal detection during 2020–2024. The total number of detected events in each region is indicated in the upper-left corner of each panel.

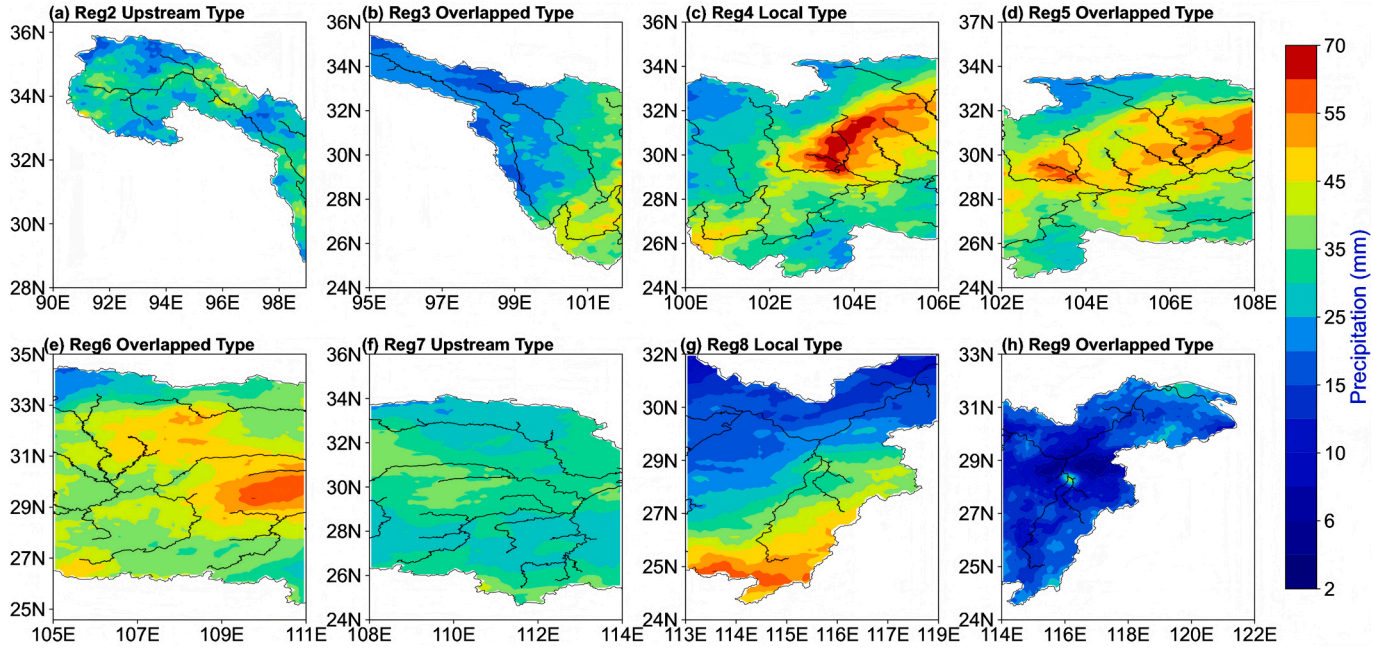


Fig. 8. Spatial distribution of cumulative precipitation within five days following DetRF-detected atmospheric circulation signals across the YRB subregions during 2020–2024.

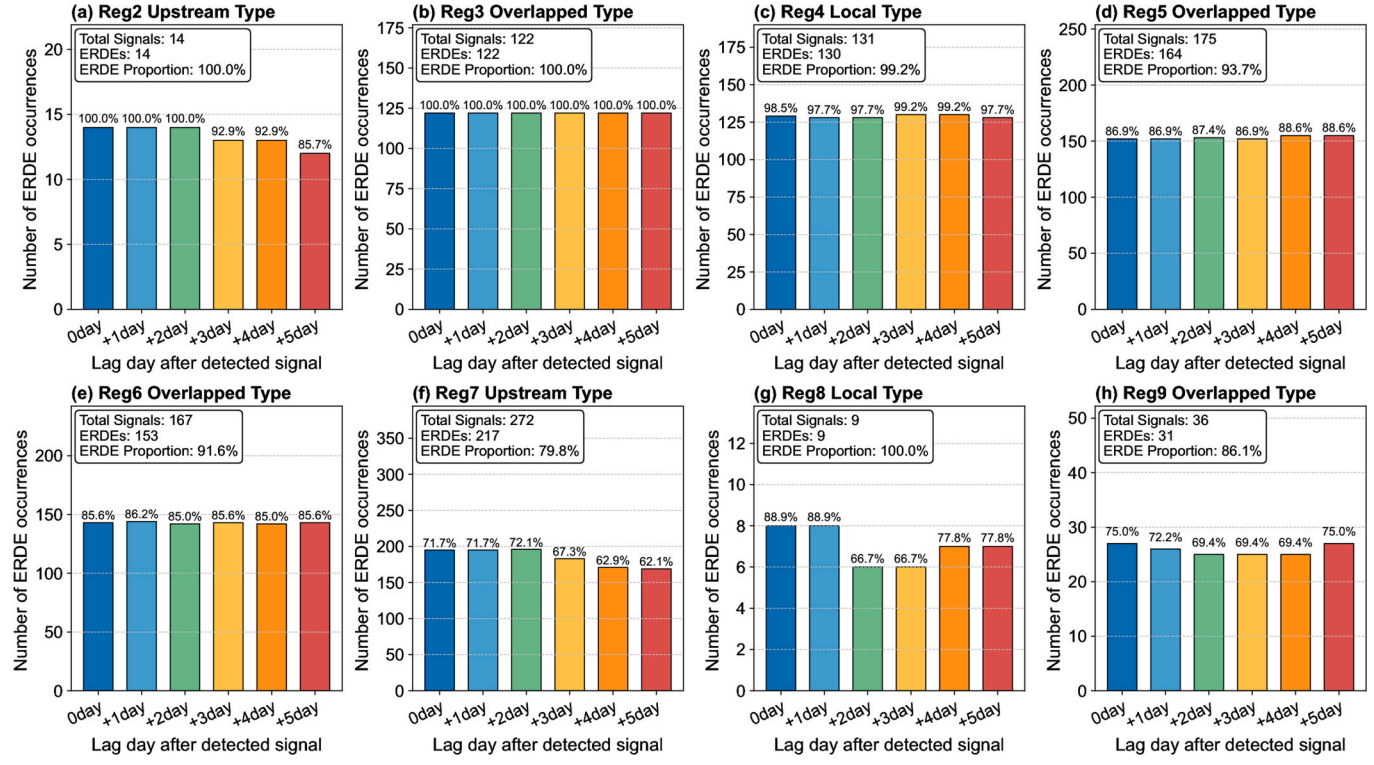


Fig. 9. Number and proportion of ERDE occurrences within five days following DetRF-detected atmospheric circulation signals. The upper-left corner shows the total number of detected signals, the number of ERDE occurrences within five days after the signal, and the corresponding proportion. The histograms illustrate the daily proportion of ERDE occurrences from day 0 to day +5 relative to the total number of detected signals during March–November of 2020–2024.

comparative analysis using six representative benchmark models across all eight YRB subregions. The selected models included extreme gradient boosting (XGBoost), HistGradientBoosting (HistGB), extremely randomized trees (ExtraTrees), convolutional neural network (CNN), isolation forest (Det), and random forest (RF) (He et al., 2016; Docheshmeh Gorgij et al., 2022; Zhu et al., 2023; Verma et al., 2024; Thakrar et al., 2026). These models were selected because they represent several widely used approaches for rare-event classification. XGBoost and HistGB are boosting-based algorithms with strong nonlinear learning capability. ExtraTrees and RF represent ensemble tree classifiers that are well-suited for high-dimensional predictors. CNN was included to assess whether direct spatial feature extraction from

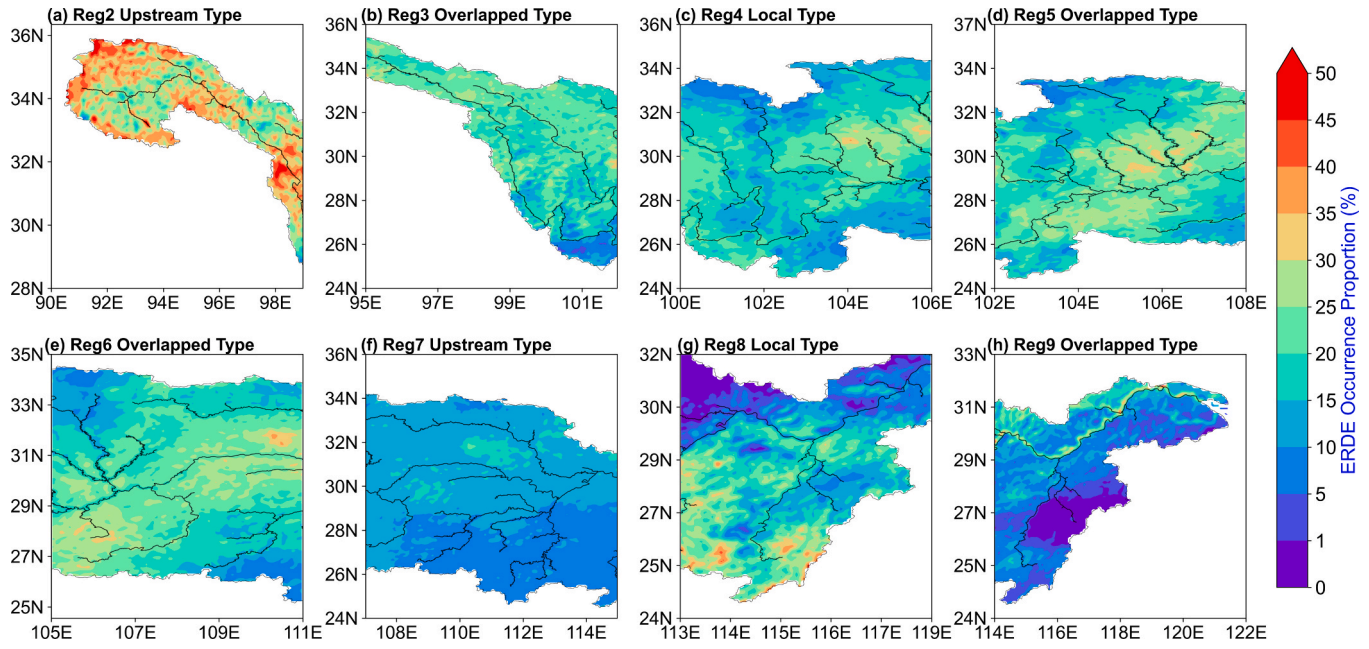


Fig. 10. Spatial distribution of ERDE occurrence proportion within five days following DetRF-detected atmospheric circulation signals during March–November of 2020–2024.

atmospheric fields could improve signal detection, whereas isolation forest was used to evaluate the effectiveness of unsupervised anomaly detection for identifying ERDE-related circulation patterns.

The benchmark results revealed substantial differences among model types. Among the baseline models, ExtraTrees achieved the best overall performance, with mean precision, recall, FNR, and F1-score across the eight subregions of 0.408, 0.619, 0.383, and 0.486, respectively (Table 4). These results suggest that randomized ensemble trees can effectively capture nonlinear relationships in high-dimensional atmospheric predictors. RF achieved a higher recall of 0.653 but a substantially lower precision of 0.274, resulting in a mean F1-score of 0.381. This indicates that RF was sensitive to potential ERDE precursor signals but generated a relatively large number of false alarms. HistGB and XGBoost exhibited comparable but slightly weaker performance, with mean F1-scores of 0.375 and 0.364, respectively. Their lower recall values suggest that gradient-boosting methods may fail to identify a portion of ERDE-related signals under strong class imbalance. The CNN model achieved a relatively high mean recall of 0.652, comparable to RF, but its precision was only 0.143, yielding a mean F1-score of 0.223. This result indicated that although CNN can extract spatial features from atmospheric fields, it tended to over-detect positive samples when the number of ERDE cases was limited. Det performed the worst, with mean precision, recall, and F1-score values of only 0.026, 0.035, and 0.028, respectively. This finding confirms that ERDE precursor signals are not merely unsupervised atmospheric outliers, but rather constitute a supervised classification problem that requires explicit event labels.

Compared with all benchmark models, DetRF achieved the most favorable balance between missed detections and false alarms. Across the eight subregions, DetRF with the year-wise temporal validation attained mean precision, recall, and F1-score values of 0.506, 0.756, and 0.576, respectively. These results demonstrated that DetRF outperformed conventional machine learning, deep learning, and unsupervised anomaly detection approaches in identifying atmospheric precursor signals associated with ERDEs in the YRB. Its advantage was reflected not only in the highest overall F1-score but also in its consistently high recall across multiple subregions, which is particularly important for the early screening of potential extreme runoff events.

Table 4
Performance comparison of six representative benchmark models across eight YRB subregions during March–November of 2000–2019.

Model	Recall	FNR	Precision	F1-score
XGBoost	0.547	0.453	0.286	0.364
HistGB	0.576	0.424	0.293	0.375
ExtraTrees	0.619	0.383	0.408	0.486
CNN	0.652	0.348	0.143	0.223
RF	0.653	0.347	0.274	0.381
Det	0.037	0.965	0.026	0.028
DetRF	**0.756**	**0.244**	**0.506**	**0.576**

4. Conclusions

This study developed the DetRF framework to identify atmospheric circulation signals associated with ERDEs across the YRB and evaluated the predictability of ERDEs within five days after signal detection using satellite-based precipitation and reanalysis river discharge data. The results establish a direct linkage between large-scale atmospheric circulation signals and subsequent hydrological extremes, demonstrating the potential of the proposed approach for ERDE early warning. The main conclusions are as follows:

(1) The DetRF framework exhibited robust and consistent performance in detecting atmospheric circulation signals associated with ERDEs across YRB. Under the more conservative year-wise temporal validation, DetRF achieved mean recall, FNR, precision, and F1-score values of 0.756, 0.245, 0.506, and 0.576, respectively. Compared with six benchmark models, DetRF provided the best overall balance between missed detections and false alarms. During the independent application period, the detected circulation patterns successfully captured key synoptic features—including the positions and intensities of troughs, vortices, and cyclones, as well as moisture transport pathways—and closely matched predefined atmospheric circulation signals.

(2) DetRF effectively captured diverse precipitation events and the spatial distribution of cumulative precipitation associated with different ERDE types. Within five days after atmospheric circulation signal detection, precipitation occurred following more

than 99% of detected signals, while persistent precipitation events accounted for over 67% of all cases across local and upstream regions. In addition, the detected signals were associated with distinct spatial precipitation patterns consistent with the underlying physical mechanisms of different ERDE types. Upstream-type ERDEs exhibited precipitation centers concentrated in upstream source areas, local-type ERDEs showed maxima within the target regions, and overlapped-type ERDEs presented multiple precipitation centers spanning upstream, local, and transitional zones.

(3) The detected atmospheric circulation signals demonstrated strong predictive value for ERDE occurrence across all eight YRB subregions. Within five days of signal detection, the proportion of ERDE occurrences ranged from 79.8% (Reg7) to 100% (Reg2, Reg3, and Reg8). The spatial distribution of ERDE occurrence frequency closely corresponded to cumulative precipitation patterns, providing independent hydrological evidence that the detected atmospheric circulation signals represent physically meaningful precursors of ERDE development. The temporal and spatial analyses demonstrate that DetRF effectively identifies atmospheric circulation signals that precede ERDE occurrence across the YRB.

Although the DetRF framework is transferable in principle, the atmospheric circulation labels should be redefined for each application basin according to its regional climatic background, dominant synoptic systems, topographic characteristics, precipitation-generating mechanisms, and hydrological response processes. The atmospheric variables required by DetRF are routinely available from operational numerical weather prediction systems. Future studies should assess the effects of forecast uncertainty and lead time by applying DetRF directly to operational forecast fields. In addition, ERDE development arises from complex interactions among atmospheric forcing, soil moisture conditions, land-surface properties, and watershed hydrological processes, and the complete physical chain linking atmospheric circulation to extreme river discharge remains insufficiently understood. Future research should therefore focus on elucidating the mechanisms of runoff generation and concentration to improve the process-based understanding of ERDE formation. Moreover, uncertainties inherited from ERA5 reanalysis data remain a limitation of the present study. Future work should quantify these uncertainties through systematic comparisons among multiple reanalysis products, such as ERA5, JRA-55, MERRA-2, and CRA-40.

CRediT authorship contribution statement

Xiaoke Xu: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yong Wang:** Formal analysis, Data curation. **Anning Huang:** Writing – review & editing, Formal analysis. **Yinghong Jing:** Writing – review & editing, Methodology. **Chunlei Gu:** Writing – review & editing. **Xiaojun She:** Writing – review & editing. **Lifu Zhang:** Writing – review & editing. **Yao Li:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Xiaoke Xu reports financial support was provided by National Natural Science Foundation of China. Xiaoke Xu reports a relationship with National Natural Science Foundation of China that includes: funding grants. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (42505025, 42571446, and 42501416), the Open Fund of State Key Laboratory of Remote Sensing and Digital Earth (OFSLRSS202518), the National Open Fund for Rainstorm Research (BYKJ2024Q02), and the Chongqing Municipal Science and Technology Bureau (CSTB2024YCJH-KYXM0054 and cstc2024ycjh-bgzxm0043).

Appendix

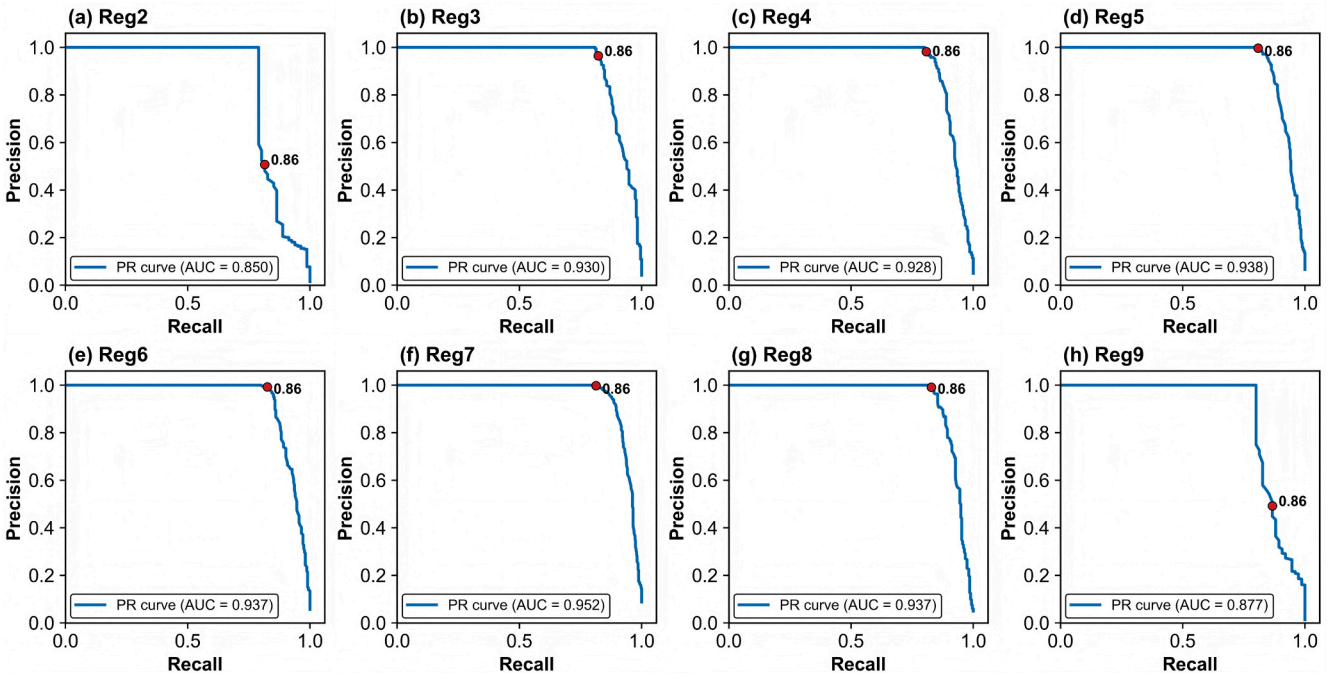


Fig. A1. Precision-recall (PR) curves of the DetRF model across the eight YRB subregions during March–November of 2000–2019. The area under the curve (AUC) provides a threshold-independent measure of classification performance, with larger AUC values indicating a better balance between precision and recall. The red dot denotes the selected probability threshold of 0.86, which was adopted for final signal classification.

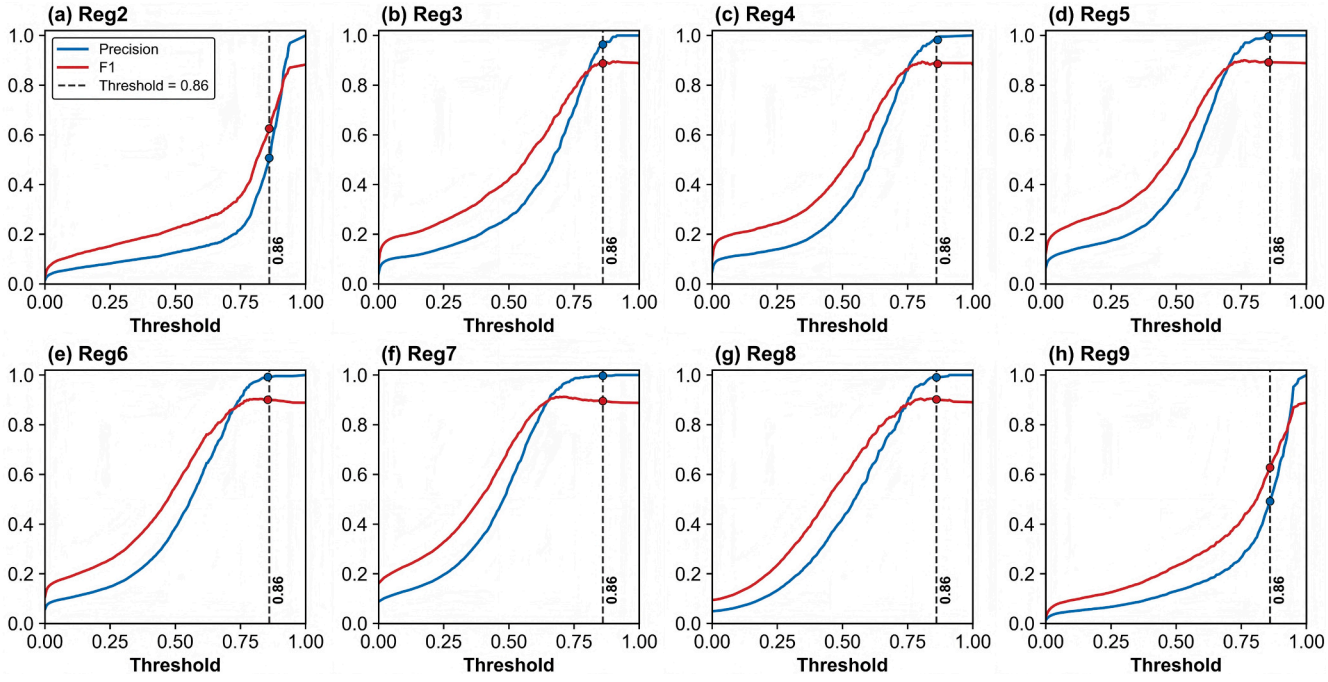


Fig. A2. Threshold-dependent precision and F1-score of the DetRF model across the eight YRB subregions during March–November of 2000–2019. The vertical dashed line indicates the selected probability threshold of 0.86. Blue and red markers represent the corresponding precision and F1-score values at the selected threshold, respectively.

Table A1
Spatial Pearson correlation coefficients between the predefined atmospheric circulation signals (2000–2019) and DetRF-detected signals (2020–2024) across the YRB. All correlations are statistically significant at the 95% confidence level ($p < 0.05$).

Regions	Pressure level (hPa)	Atmospheric variables		
		Zonal wind	Meridional wind	Geopotential heights
Reg2	200	0.981	0.967	0.959
	500	0.702	0.894	0.957
Reg3	200	0.995	0.942	0.989
	500	0.950	0.947	0.987
Reg4	600	0.945	0.887	0.979
	700	0.840	0.925	0.980
Reg5	500	0.982	0.910	0.992
	850	0.875	0.970	0.952
Reg6	500	0.963	0.927	0.990
	850	0.923	0.967	0.962
Reg7	600	0.873	0.867	0.942
	700	0.904	0.860	0.913
Reg8	850	0.723	0.739	0.836
Reg9	500	0.988	0.790	0.965
	750	0.820	0.748	0.971

Table A2
Statistics of ERDE occurrences identified from GloFAS data and detected by the DetRF model across individual subregions during the independent application period (2020–2024). ERDE occurrences separated by six days or less were grouped as a single event.

Regions	Total ERDE occurrences	DetRF-detected signals	Signals followed by an ERDE within 0–5 d	ERDE recall
Reg2	37	14	14	37.84%
Reg3	153	122	122	79.74%
Reg4	194	131	130	67.01%
Reg5	206	175	164	79.61%
Reg6	197	167	153	77.66%
Reg7	244	272	217	88.93%
Reg8	16	9	9	56.25%
Reg9	36	36	31	86.11%

Data availability

The datasets used in this study are described as follows. River order and slope information (Yamazaki et al., 2019) were obtained from the MERIT_Hydro_v07_Basins_v01 dataset (https://www.reachhydro.org/home/params/merit-basins). High-resolution daily river discharge reanalysis data (GloFAS) were obtained from the Joint Research Centre (JRC) of the European Commission (Harrigan et al., 2020) (https://data.jrc.ec.europa.eu/dataset/653b7f9b-0768-47f1-ac0a-facc3b21043e). Precipitation data (Hou et al., 2014) were sourced from the final run of the GPM-IMERG V06 dataset (https://gpm.nasa.gov/data/imerg). Atmospheric variables were taken from the fifth-generation ECWMF reanalysis dataset (ERA5) (Hersbach et al., 2023), accessed through the Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (https://doi.org/10.24381/cds.bd0915c6). Topography information (Jarvis et al., 2008) was derived from the Shuttle Radar Topography Mission (SRTM) Digital Elevation Database v4.1 (https://developers.google.com/earth-engine/datasets/catalog/CGIAR_SRTM90_V4).

References

Adikari, K.E., Shrestha, S., Ratnayake, D.T., Budhathoki, A., Mohanasundaram, S., Dailey, M.N., 2021. Evaluation of artificial intelligence models for flood and drought forecasting in arid and tropical regions. Environ Model Softw. 144, 105136. https://doi.org/10.1016/j.envsoft.2021.105136.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., Tian, Q., 2023. Accurate medium-range global weather forecasting with 3D neural networks. Nature 619, 533–538. https://doi.org/10.1038/s41586-023-06185-3.

Bližňák, V., Kašpar, M., Müller, M., Zacharov, P., 2019. Sub-daily temporal reconstruction of extreme precipitation events using NWP model simulations. Atmos. Res. 224, 65–80. https://doi.org/10.1016/j.atmosres.2019.03.019.

Breinl, K., Lun, D., Müller-Thomy, H., Blöschl, G., 2021. Understanding the relationship between rainfall and flood probabilities through combined intensity-duration-frequency analysis. J. Hydrol. 602, 126759. https://doi.org/10.1016/j.jhydrol.2021.126759.

Camps-Valls, G., Fernández-Torres, M.Á., Cohrs, K.H., Höhl, A., Castelletti, A., Pacal, A., 2025. Artificial intelligence for modeling and understanding extreme weather and climate events. Nat. Commun. 16, 1919. https://doi.org/10.1038/s41467-025-56573-8.

Charlton-Perez, A.J., Dacre, H.F., Driscoll, S., Gray, S.L., Harvey, B., Harvey, N.J., 2024. Do AI models produce better weather forecasts than physics-based models? A quantitative evaluation case study of storm Ciarán. npj Clim. Atmos. Sci. 7, 93. https://doi.org/10.1038/s41612-024-00638-w.

Cui, L., He, M., Zou, Z., Yao, C., Wang, S., An, J., Wang, X., 2022. The influence of climate change on droughts and floods in the Yangtze River basin from 2003 to 2020. Sensors 22, 8178. https://doi.org/10.3390/s22218178.

Cui, L., Meng, J., Zhong, B., An, J., Li, Y., Guo, H., Lu, Y., 2025. Spatiotemporal evolution of 1998 extreme flood event in the Yangtze River basin from the perspective of the reconstructed GRACE/GRACE-FO data. J. Hydrol.: Reg. Stud. 60, 102551. https://doi.org/10.1016/j.ejrh.2025.102551.

de Sousa Araújo, A., Silva, A.R., Zárate, L.E., 2022. Extreme precipitation prediction based on neural network model-a case study for southeastern Brazil. J. Hydrol. 606, 127454. https://doi.org/10.1016/j.jhydrol.2022.127454.

Davenport, F.V., Diffenbaugh, N.S., 2021. Using machine learning to analyze physical causes of climate change: a case study of US Midwest extreme precipitation. Geophys. Res. Lett. 48, e2021GL093787. https://doi.org/10.1029/2021GL093787.

Deng, P., Zhang, M., Hu, Q., Wang, L., Bing, J., 2022. Pattern of spatio-temporal variability of extreme precipitation and flood-waterlogging process in Hanjiang River basin. Atmos. Res. 276, 106258. https://doi.org/10.1016/j.atmosres.2022.106258.

Ding, F., Li, C., 2017. Subtropical westerly jet waveguide and winter persistent heavy rainfall in south China. J. Geophys. Res. Atmos. 122, 7385–7400. https://doi.org/10.1002/2017JD026530.

Ding, Z., Ha, Y., Zhong, Z., 2024. Summer extreme precipitation patterns and synoptic-scale circulation precursors over the Tibetan Plateau. Sci. China Earth Sci. 67, 1625–1638. https://doi.org/10.1007/s11430-023-1321-6.
Docheshmeh Gorgij, A., Alizamir, M., Kisi, O., Elshafie, A., 2022. Drought modelling by standard precipitation index (SPI) in a semi-arid climate using deep learning method: long short-term memory. Neural Comput. Appl. 34, 2425–2442. https://doi.org/10.1007/s00521-021-06505-6.
Fang, J., Kong, F., Fang, J., Zhao, L., 2018. Observed changes in hydrological extremes and flood disaster in Yangtze River Basin: spatial-temporal variability and climate change impacts. Nat. Hazards 93, 89–107. https://doi.org/10.1007/s11069-018-3290-3.
Frame, J.M., Kratzert, F., Klotz, D., Gauch, M., Shalev, G., Gilon, O., 2022. Deep learning rainfall-runoff predictions of extreme events. Hydrol. Earth Syst. Sci. 26, 3377–3392. https://doi.org/10.5194/hess-26-3377-2022.
Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., et al., 2020. GloFAS-ERA5 operational global river discharge reanalysis 1979–present. Earth Syst. Sci. Data 12, 2043–2060. https://doi.org/10.5194/essd-12-2043-2020.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. Springer, New York. https://doi.org/10.1007/978-0-387-84858-7.
He, X., Chaney, N.W., Schleiss, M., Sheffield, J., 2016. Spatial downscaling of precipitation using adaptable random forests. Water Resour. Res. 52, 8217–8237. https://doi.org/10.1002/2016WR019034.
Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al., 2023. ERA5 hourly data on pressure levels from 1940 to present [dataset]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.bd0915c6.
Hodnebrog, Ø., Steensen, B.M., Marelle, L., Alterskjær, K., Dalsøren, S.B., Myhre, G., 2022. Understanding model diversity in future precipitation projections for South America. Clim. Dyn. 58, 1329–1347. https://doi.org/10.1007/s00382-021-05964-w.
Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., 2014. The global precipitation measurement mission. Bull. Am. Meteorol. Soc. 95, 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1.
Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. The Shuttle radar topography mission (SRTM) 90 m digital elevation database v4.1. CGIAR Consortium for Spatial Information. https://developers.google.com/earth-engine/datasets/catalog/CGIAR_SRTM90_V4.
Jia, H., Chen, F., Pan, D., Du, E., Wang, L., Wang, N., Yang, A., 2022. Flood risk management in the Yangtze River basin-comparison of 1998 and 2020 events. Int. J. Disaster Risk Reduct. 68, 102724. https://doi.org/10.1016/j.ijdrr.2021.102724.
Kafi, K.M., Ponrahono, Z., 2024. Advances in weather and climate extreme studies: a systematic comparative review. Discov. Geosci. 2, 66. https://doi.org/10.1007/s44288-024-00079-1.
Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proc. 14th Int. Jt. Conf. Artif. Intell. (IJCAI'95), Montreal, Canada, pp. 1137–1143.
Kumar, V., Sharma, K.V., Caloiero, T., Mehta, D.J., Singh, K., 2023. Comprehensive overview of flood modeling approaches: a review of recent advances. Hydrology 10, 141. https://doi.org/10.3390/hydrology10070141.
Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., et al., 2023. Learning skillful medium-range global weather forecasting. Science 382, 1416–1421. https://doi.org/10.1126/science.adi2336.
Li, L., Zhang, R., 2023. Evolution mechanisms, impacts, and variations of the vortices originated from the Tibetan Plateau. Earth Sci. Rev. 242, 104463. https://doi.org/10.1016/j.earscirev.2023.104463.
Liu, C., Guo, L., Ye, L., Zhang, S., Zhao, Y., Song, T., 2018. A review of advances in China's flash flood early-warning system. Nat. Hazards 92, 619–634. https://doi.org/10.1007/s11069-018-3173-7.
Liu, F.T., Ting, K.M., Zhou, Z.H., 2008. Isolation forest. In: Proceedings of the 2008 Eighth IEEE International Conference on Data Mining (ICDM 2008). IEEE, Pisa, Italy, pp. 413–422. https://doi.org/10.1109/ICDM.2008.17.
Liu, X., Zhang, Y., He, B., Liu, Y., Wu, G., Xu, J., 2025. The month-to-year precursory and synchronous inherent connections between global oceanic modes and extreme precipitation over China. Adv. Atmos. Sci. 42, 1513–1532. https://doi.org/10.1007/s00376-024-4306-4.
Mukkavilli, S.K., Civitarese, D.S., Schmude, J., Jakubik, J., Jones, A., Nguyen, N., 2023. AI foundation models for weather and climate: applications, design, and implementation. arXiv:2309.10808. https://arxiv.org/abs/2309.10808. (Accessed 5 September 2026).
Nearing, G., Cohen, D., Dube, V., Gauch, M., Gilon, O., Harrigan, S., et al., 2024. Global prediction of extreme floods in ungauged watersheds. Nature 627, 559–563. https://doi.org/10.1038/s41586-024-07145-1.
Nie, Y., Sun, J., 2021. Synoptic-scale circulation precursors of extreme precipitation events over southwest China during the rainy season. J. Geophys. Res. Atmos. 126, e2021JD035134. https://doi.org/10.1029/2021JD035134.
Prein, A.F., Rasmussen, R.M., Ikeda, K., Liu, C., Clark, M.P., Holland, G.J., 2017. The future intensification of hourly precipitation extremes. Nat. Clim. Chang. 7, 48–52. https://doi.org/10.1038/nclimate3168.
Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T.R., El-Kadi, A., Masters, D., et al., 2025. Probabilistic weather forecasting with machine learning. Nature 637, 84–90. https://doi.org/10.1038/s41586-024-08252-9.
Provost, F., Fawcett, T., 2001. Robust classification for imprecise environments. Mach. Learn. 42 (3), 203–231. https://doi.org/10.1023/A:1007601015854.
Rawal, U., Patel, S., 2025. Anomaly detection in meteorological data using machine learning techniques. In: Proceedings of the 2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS). IEEE, pp. 1–6. https://doi.org/10.1109/SCEECS64059.2025.10940145.
Sillmann, J., Thorarinsdottir, T., Keenlyside, N., Schaller, N., Alexander, L.V., Hegerl, G., Zwiers, F.W., 2017. Understanding, modeling and predicting weather and climate extremes: challenges and opportunities. Weather Clim. Extrem. 18, 65–74. https://doi.org/10.1016/j.wace.2017.10.003.
Tang, J., Lu, Y., Wang, S., Guo, Z., Lu, Y., Fang, J., 2023. Projection of hourly extreme precipitation using the WRF model over eastern China. J. Geophys. Res. Atmos. 128, e2022JD036448. https://doi.org/10.1029/2022JD036448.
Thakrar, C., Gachancipa, N., Deshpande, K., Bals, A.M., Paxton, L., 2026. Categorizing characteristic regions of high-latitude scintillations using a combination of isolation forest and neural network machine learning algorithms. J. Geophys. Res. Mach. Learn. Comput. 3, e2024JH000298. https://doi.org/10.1029/2024JH000298.
Try, S., Tanaka, S., Tanaka, K., Sayama, T., Khujanazarov, T., Oeurng, C., 2022. Comparison of CMIP5 and CMIP6 GCM performance for flood projections in the Mekong River Basin. J. Hydrol.: Reg. Stud. 40, 101035. https://doi.org/10.1016/j.ejrh.2022.101035.
Tuel, A., Schaefli, B., Zscheischler, J., Martius, O., 2022. On the links between sub-seasonal clustering of extreme precipitation and high discharge in Switzerland and Europe. Hydrol. Earth Syst. Sci. 26, 2649–2672. https://doi.org/10.5194/hess-26-2649-2022.
Verma, P., Verma, R., Mallet, M., Sisodiya, S., Zare, A., Dwivedi, G., Ristovski, Z., 2024. Assessment of human and meteorological influences on PM10 concentrations: insights from machine learning algorithms. Atmos. Pollut. Res. 15, 102123. https://doi.org/10.1016/j.apr.2024.102123.
Wang, B., Wei, W., Yin, Z., Xu, L., 2024. Using machine learning to analyze the changes in extreme precipitation in southern China. Atmos. Res. 302, 107307. https://doi.org/10.1016/j.atmosres.2024.107307.
Wang, Q., Xu, Y., Cai, X., Tang, J., Yang, L., 2021. 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.
Wang, X., Xia, J., Zhou, M., Deng, S., Li, Q., 2022. Assessment of the joint impact of rainfall and river water level on urban flooding in Wuhan City, China. J. Hydrol. 613, 128419. https://doi.org/10.1016/j.jhydrol.2022.128419.
Wei, K., Ouyang, C., Duan, H., Li, Y., Chen, M., Ma, J., Zhou, S., 2020. Reflections on the catastrophic 2020 Yangtze River Basin flooding in southern China. Innovation 1, 100038. https://doi.org/10.1016/j.xinn.2020.100038.
Wright, D.B., Yu, G., England, J.F., 2020. Six decades of rainfall and flood frequency analysis using stochastic storm transposition: review, progress, and prospects. J. Hydrol. 585, 124816. https://doi.org/10.1016/j.jhydrol.2020.124816.
Xie, J., Xu, Y.P., Yu, H., Huang, Y., Guo, Y., 2022. Monitoring the extreme flood events in the Yangtze River basin based on GRACE and GRACE-FO satellite data. Hydrol. Earth Syst. Sci. 26, 5933–5954. https://doi.org/10.5194/hess-26-5933-2022.
Xie, T., Huang, W., Chang, S., Zheng, F., Chen, J., Chen, J., Chen, F., 2021. Moisture sources of extreme precipitation events in arid Central Asia and their relationship with atmospheric circulation. Int. J. Climatol. 41, E271–E282. https://doi.org/10.1002/joc.6683.
Xie, Y., Li, X., Ngai, E.W.T., Ying, W., 2009. Customer churn prediction using improved balanced random forests. Expert Syst. Appl. 36, 5445–5449. https://doi.org/10.1016/j.eswa.2008.06.121.
Xu, Q., Chen, J., Peart, M.R., Ng, C.N., Hau, B.C., Law, W.W., 2018. Exploration of severities of rainfall and runoff extremes in ungauged catchments: a case study of Lai Chi Wo in Hong Kong, China. Sci. Total Environ. 634, 640–649. https://doi.org/10.1016/j.scitotenv.2018.04.024.
Xu, X., Huang, A., Jing, Y., Gu, C., She, X., Li, Y., 2026. Escalating extreme river discharge events driven by precipitation changes in the Yangtze River Basin. Earths Future 14, e2025EF007133. https://doi.org/10.1029/2025EF007133.
Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P.D., Allen, G.H., Pavelsky, T.M., 2019. MERIT Hydro: a high-resolution global hydrography map based on latest topography dataset. Water Resour. Res. 55, 5053–5073. https://doi.org/10.1029/2019WR024873.
Yang, P., Xia, J., Luo, X., Meng, L., Zhang, S., Cai, W., Wang, W., 2021. Impacts of climate change-related flood events in the Yangtze River Basin based on multi-source data. Atmos. Res. 263, 105819. https://doi.org/10.1016/j.atmosres.2021.105819.
Yang, P., Zhang, S., Xia, J., Zhan, C., Cai, W., Wang, W., Li, J., 2022. Analysis of drought and flood alternation and its driving factors in the Yangtze River Basin under climate change. Atmos. Res. 270, 106087. https://doi.org/10.1016/j.atmosres.2022.106087.
Yin, J., Gentine, P., Zhou, S., Sullivan, S.C., Wang, R., Zhang, Y., Guo, S., 2018. Large increase in global storm runoff extremes driven by climate and anthropogenic changes. Nat. Commun. 9, 4389. https://doi.org/10.1038/s41467-018-06765-2.
Yin, J., Guo, S., Gentine, P., Sullivan, S.C., Gu, L., He, S., Liu, P., 2021. Does the hook structure constrain future flood intensification under anthropogenic climate warming? Water Resour. Philos. Phenomenol. Res. 57, e2020WR028491. https://doi.org/10.1029/2020WR028491.
Zeng, Z., Wang, D., 2022. Observations of heavy short-term rainfall hotspots associated with warm-sector episodes over coastal South China. Atmos. Res. 276, 106273. https://doi.org/10.1016/j.atmosres.2022.106273.
Zhang, L., Liu, Y., Ren, L., Teuling, A.J., Zhu, Y., Wei, L., Yin, H., 2022. Analysis of flash droughts in China using machine learning. Hydrol. Earth Syst. Sci. 26, 3241–3261. https://doi.org/10.5194/hess-26-3241-2022.

Zhao, G., Liu, R., Yang, M., Tu, T., Ma, M., Hong, Y., Wang, X., 2022. Large-scale flash flood warning in China using deep learning. J. Hydrol. 604, 127222. https://doi.org/10.1016/j.jhydrol.2021.127222.

Zhu, H., Liu, H., Zhou, Q., Cui, A., 2023. A XGBoost-based downscaling-calibration scheme for extreme precipitation events. IEEE Trans. Geosci. Remote Sens. 61, 1–12. https://doi.org/10.1109/TGRS.2023.3294266.

Zhu, R., Chen, L., 2023. Climatic characteristics of the Jianghuai cyclone and its linkage with precipitation during the Meiyu period from 1961 to 2020. EGUspher. https://doi.org/10.5194/egusphere-2023-2008.