

Large-scale flood monitoring based on time-series Sentinel-1 images and Z-score

Runmei Xing^{1,2}, Fengyuan Lu³, Jun Zhao^{1*}

¹ School of Civil Engineering, Zhengzhou University, Zhengzhou 450001, China

5 ² Henan College of Transportation, Zhengzhou 451460, China

³ School of Computer Science and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China

Correspondence to: Jun Zhao (zhaoj@zzu.edu.cn)

Abstract. Synthetic Aperture Radar (SAR) satellites have emerged as a crucial technology for real-time flood monitoring in the face of escalating flood disaster hazards. However, existing flood extraction techniques still face challenges in simultaneously achieving accurate urban inundation detection and extensive large-scale monitoring. To address this challenge, this study proposes a time-series Sentinel-1 SAR anomaly-detection approach based on an analysis of backscatter characteristics across different land-cover types and the integration of auxiliary datasets. First, an analysis of scattering behaviour in time-series SAR images revealed that, in natural areas, the consistency of backscatter intensity is strongly influenced by vegetation growth status. In building areas, rainfall and floodwater can intensify or suppress double-bounce scattering depending on building geometry and water depth, thereby disrupting intensity consistency. Based on these findings, a Z-score-based flood classification tree was developed. The method uses reference images and flood-period images to compute standardized anomaly maps with statistical interpretability under approximate normality and establishes operationally transparent flood-detection thresholds. The integration of VV and VH polarizations within the classification tree further improves the reliability of flood identification. Applied to the 2021 Weihui flood event, the method demonstrated strong performance, achieving a critical success index (CSI) of 60% and overall accuracy (OA) of 90% in natural areas, and a CSI of 62% and OA of 73% in building areas. The proposed approach shows significant advantages in accurately classifying flood-affected areas and offers the capability to monitor both large-scale floods and urban inundation.

1 Introduction

Flooding is one of the most frequent, widespread, and economically devastating natural disasters worldwide (Jongman et al., 2012; Barredo, 2007; Tingsanchali, 2012). Statistics show that between 2000 and 2019, more than 1.6 billion people globally were affected by floods to varying degrees, with Asia alone accounting for 1.5 billion of those cases (Tellman et al., 2021). As global climate change intensifies, extreme weather events are becoming more frequent, and both the occurrence and severity of floods are on the rise. Satellite-based studies indicate that from 2000 to 2015, the proportion of the global population affected by floods increased by more than 20% (Voigt et al., 2016). It is projected that by 2030, the proportion of people affected by flooding in Asia and Africa will continue to increase (Voigt et al., 2016).

Extreme rainfall is the primary driver of widespread flooding (Kunkel, 2003). However, due to cloud cover, traditional optical and infrared remote sensing satellites often struggle to capture the actual conditions on the ground, making it difficult to promptly map the full extent of flood events and frequently leading to underestimation of the resulting damage (Anusha and Bharathi, 2020; Zhang et al., 2002; Alvarez-Vanhard et al., 2021). In contrast, synthetic aperture radar (SAR) is an active microwave sensing technology (Moreira et al., 2013; Cruz et al., 2022), enabling it to acquire surface information under all weather conditions and providing all-day, all-weather imaging capabilities (Cruz et al., 2022; Liu et al., 2025). Synthetic aperture radar satellites are a key component of the global Earth observation satellite system and play an indispensable role in flood emergency response. The European Space Agency's Sentinel-1, for instance, was deployed shortly after launch to support emergency response during flood events in Namibia and Bosnia and Herzegovina (Nhangumbe et al., 2023; Twele et al., 2016; Martinis et al., 2018). Similarly, China's Gaofen-3 captured multiple emergency images during the catastrophic "7•20" Zhengzhou rainstorm. These cases demonstrate that SAR has become one of the preferred tools for flood monitoring (Kang et al., 2018; Zhao et al., 2021; Wang et al., 2025). Moreover, the rapid development of commercial SAR satellites has introduced data with varying spatial resolutions, frequency bands, and polarization modes, significantly enhancing SAR's ability to map dynamic surface changes during emergency responses (Zhao et al., 2021; Martinis et al., 2015; Amitrano et al., 2024; Zafarmomen et al., 2026).

The timeliness and accuracy of flood mapping are critical for the effective allocation of rescue resources and the reduction of casualties. In natural areas, floods are generally easy to identify using SAR images, as the smooth surface of water causes strong specular reflection, leading to a significant decrease in backscatter values (Tripathi et al., 2020). Medium-resolution SAR (MR-SAR) images are often sufficient for effective monitoring in such cases. However, in building areas, the scattering characteristics are more complex. Backscatter values are influenced by factors such as the satellite orbit (ascending or descending), polarization mode, and the orientation and distribution of buildings (Wang and Tong, 2018). As a result, high-resolution SAR images are typically required for effective flood detection in urban environments (Cian et al., 2018; Voigt et al., 2009). The wide variability in the scale of flood disasters poses additional challenges for monitoring methods and technologies. Although high-resolution SAR images provide greater detail, their narrower swaths limit their use in large-scale flood monitoring (Liu et al., 2020). At the same time, traditional MR-SAR image monitoring methods may exhibit significant errors in densely populated building areas, leading to more severe consequences. Therefore, it is of great significance to develop methods that can support both large-scale monitoring and the detection of urban flooding simultaneously.

To address this issue, two approaches exist. The first involves expanding the coverage of high-resolution SAR imagery, but this depends on satellite availability and scheduling, making it difficult to achieve at present. Therefore, we opted for the second approach: optimizing methods based on MR-SAR imagery to ensure rapid, large-scale flood monitoring and effective detection of urban flooding in built-up areas. This study integrates the characteristics of SAR backscatter with time-series anomaly detection concepts to improve both the accuracy and timeliness of inundation detection in urban environments. To address the challenges posed by complex surface cover types and high heterogeneity in building areas, this study analyzes the scattering mechanisms in building regions based on variations in microwave scattering features. Recognizing the distinct flood

65 response behaviors in natural and building areas, we propose a time series SAR-based anomaly flood monitoring method. By incorporating auxiliary data and dual-polarization SAR information, this approach achieves high robustness and accuracy in flood detection across both natural and urban environments.

Compared with conventional single-image thresholding, image differencing, or single-polarization SAR flood mapping, the contribution of this study is threefold. First, pixel-wise temporal standardization is used to reduce dependence on event-specific
70 global thresholds and to improve transferability across heterogeneous natural and built-up surfaces. Second, VV and VH anomalies are integrated in a rule-based classification tree, allowing rough open water, vegetation-related changes, and urban double-bounce effects to be considered within a unified framework. Third, the method remains operationally scalable because it only requires Sentinel-1 GRD time series, auxiliary water-occurrence data, and simple statistical operations; therefore, it can be implemented for rapid large-area flood response while retaining sensitivity to urban inundation.

75 **2 Study area and data**

2.1 Study area

Weihui City is located in the northern part of Henan Province, China, covering a total area of 862 km², with a north-south length of 43 km and an east-west width of 35 km (Fig. 1). Situated within the Hai River Basin, the city lies at the confluence of multiple rivers, including the Wei River and the Communism Canal, resulting in a large surface water area. The urban area
80 is built on a plain formed by river sedimentation, featuring low-lying terrain that slopes from east to west, which hinders effective drainage of accumulated water. In recent years, rapid urbanization has altered the characteristics of the underlying surface, further increasing the risk of urban waterlogging. Climatically, Weihui experiences a temperate continental monsoon climate, characterized by hot and rainy summers. From July to August, precipitation accounts for approximately 50% of the annual total, often concentrated in a few intense rainstorms. Moreover, interannual variation in precipitation is significant—
85 wet years may see rainfall as high as 1000 mm, while dry years may receive only 400 mm, a difference of more than 2.5 times. This results in highly variable river flows within the city, making summer the peak season for flood disasters in Weihui (Du et al., 2024; Zhang et al., 2022).

Starting on July 17, 2021, persistent heavy rainfall triggered widespread flooding across many parts of Henan Province. Records from the Xinxiang Meteorological Station indicated that this extraordinary rainstorm was the most intense since
90 observations began, with a maximum recorded rainfall of 965.5 mm. Located downstream, Weihui City faced enormous pressure from upstream flood discharge even after the rainfall had ceased. River water levels continued to rise, eventually surpassing the warning thresholds. Overflows from the Wei River and the Communism Canal caused massive volumes of floodwater to inundate the urban area of Weihui, leading to extremely severe urban flooding. At its peak, 79% of the city's building area was submerged, with floodwater volume reaching 21,000,000 m³. All four reservoirs in the city overflowed,
95 making it impossible to discharge floodwaters via the river system. Compounding the issue, the city's power supply was disrupted by the flood, preventing pumping stations from operating at full capacity. As a result, floodwaters in the main urban

area were not fully drained until July 31. Nearly 300,000 people were affected, and over 14,000 houses were inundated. In addition to the urban devastation, all 13 townships under Weihui's jurisdiction experienced varying degrees of flooding, with approximately 330 km² of crops damaged (Zhang et al., 2022).

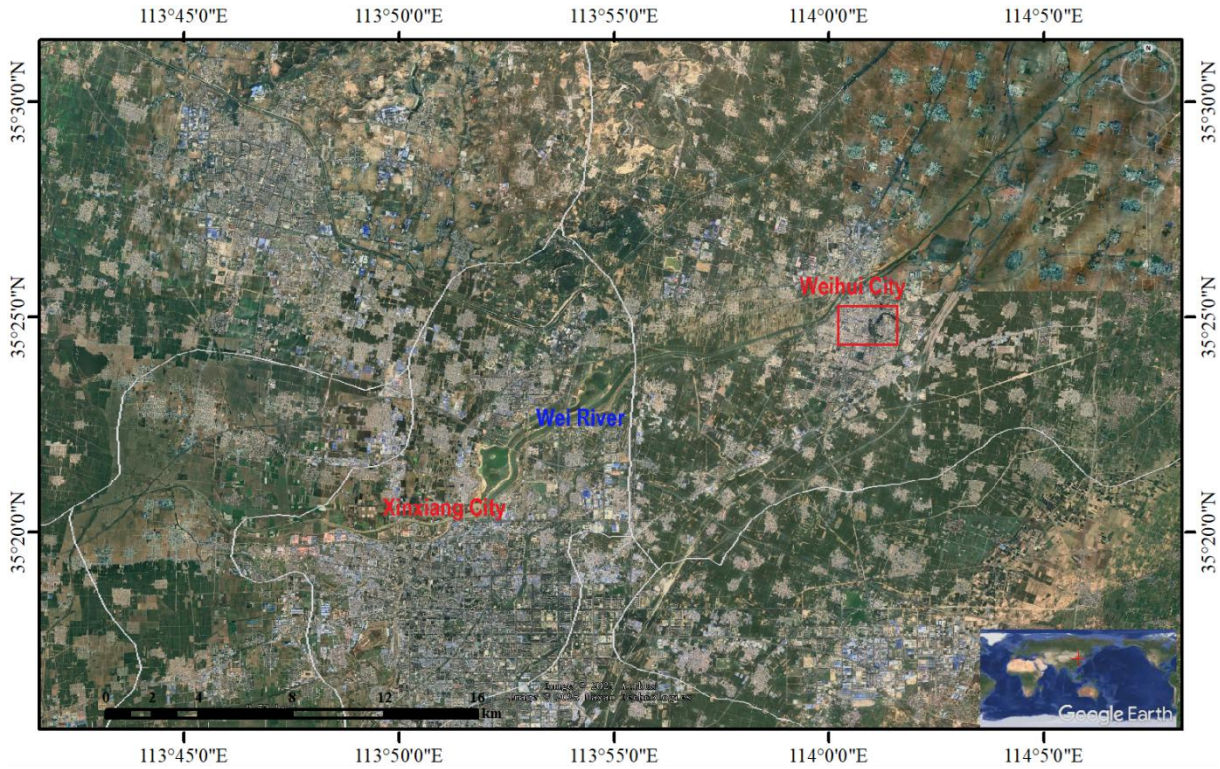


Figure 1: The geographical location of the study area (Map data © Google). The red box represents the UAV imaging coverage.

2.2 Sentinel-1 data

The Sentinel-1 satellite constellation currently consists of three satellites: Sentinel-1A, launched in 2014; Sentinel-1B, launched in 2016; and Sentinel-1C, launched in 2024. Sentinel-1B was officially decommissioned in August 2022 due to an irreparable power failure (Torres et al., 2012; Geudtner et al., 2014). When operating with a single satellite, Sentinel-1 has a revisit cycle of 12 days, with observation frequency reaching up to 6 days in certain regions (Torres et al., 2012). To meet diverse application needs, Sentinel-1 is equipped with four imaging modes: stripmap mode (SM), interferometric wide swath mode (IW), extra wide swath mode (EW), and wave mode (WV). Among these, the IW mode is the primary mode for land surface observation. It enables coverage of the vast majority of the Earth's land areas, with each acquisition providing a swath width of 250 km and a spatial resolution of 5×20 m.

This study primarily utilizes the intensity data from Sentinel-1 ground range detected (GRD) products for flood extraction (Nhangumbe et al., 2023; Torres et al., 2012). The data processing involves the following key steps: (1) Precise orbit correction – applying refined satellite orbit data to improve geolocation accuracy. (2) Thermal noise removal – eliminating sensor-induced

noise to enhance image quality. (3) Radiometric calibration (*Sigma0*) – converting pixel values into physically meaningful
115 backscatter intensity using the sensor calibration metadata embedded in the images. (4) Speckle filtering – applying Refined
Lee filtering to reduce speckle noise while preserving image details. (5) Terrain correction – using a digital elevation model
(DEM) to correct the influence of surface topography on backscatter intensity, resulting in a normalized backscatter coefficient.
(6) Conversion to decibel scale (dB) – expressing the backscatter coefficient on a logarithmic scale to facilitate easier
computation and statistical analysis, especially for values that span several orders of magnitude.
120 The Sentinel-1 images used in this study correspond to Path 40 (112) and Frame 113 (111), covering the period from January
4, 2021, to December 30, 2021. The flood event occurred in mid-to-late July, with large-scale floodwaters receding
approximately by mid-to-late August.

2.3 GPM data

The Global Precipitation Measurement (GPM) mission, developed through international collaboration among multiple space
125 agencies, enables frequent and high-accuracy global precipitation observations. The satellite is equipped with two main
instruments for precipitation monitoring: the dual-frequency precipitation radar (DPR) and the GPM Microwave Imager (GMI).
The Integrated Multi-satellite Retrievals for GPM (IMERG) precipitation product, provided by the National Aeronautics and
Space Administration (NASA), is one of the key datasets derived from GPM observations (Tang et al., 2020). This product is
primarily generated using GPM satellite data and is supplemented with data from other meteorological satellites, including
130 TRMM, METOP, and NOAA (Kummerow et al., 2000). By integrating observations across various frequency bands and
resolutions, the IMERG product enhances both the accuracy and spatiotemporal resolution of precipitation estimates. It offers
global precipitation data at a temporal resolution of 30 minutes and a spatial resolution of 0.1°. In this study, the
“precipitationCal” band from the IMERG product is used as the reference source for precipitation data. This band has been
calibrated using ground-based observation station data and is the officially recommended product in the IMERG user manual.
135 GPM data are mainly used to characterize rainfall intensity during the flood event and to assist in validating the flood extraction
results.

2.4 Auxiliary data

The Global Surface Water (GSW) dataset is a 30 m resolution product that provides global surface water distribution
information spanning the past 38 years (Pekel et al., 2016). It is developed and continuously updated by the Joint Research
140 Center (JRC). Each pixel is classified as either water or non-water using an expert system, and then temporal analysis is applied
to capture the historical dynamics of surface water presence. The algorithm proposed in this study uses the occurrence band
from the GSW dataset, which represents the frequency of water occurrence for each pixel, with values ranging from 0 to 100%.
Based on threshold values, surface water can be categorized into permanent water, seasonal water, and historic floodwater.
The DEM data used in this study are derived from the European Space Agency's Copernicus DEM, with a spatial resolution
145 of 30 m (Li et al., 2022).

3 Methods

This study first uses a year-long, same-orbit Sentinel-1 time series to test the normality of pixel-wise backscatter (via skewness and kurtosis). VV and VH images are then standardized by the mean and standard deviation to produce Z-score anomaly maps, with thresholds set considering acquisition differences. To avoid missing persistently inundated areas, GSW water-occurrence frequency is used to mask seasonal/permanent water. A decision-tree classifier assigns flood labels in natural and building areas using separate thresholds, distinguishing dual- or single-polarization floods. Finally, speckle-induced artifacts are removed by post-processing (e.g., small-patch filtering), and the results are validated with Sentinel-2/UAV imagery and assessed using binary metrics.

3.1 Outlier detection

The Z-score method is a statistical anomaly detection method that has been proven effective for monitoring newly formed surface water bodies. Compared with other anomaly detection approaches, the Z-score method relies on commonly used statistical measures—mean and standard deviation (std)—which offer strong interpretability. Thresholds can be defined based on the desired confidence level, providing solid theoretical support for better understanding the data and results (Fernandez et al., 2016; Cheadle et al., 2003). The Z-score method standardizes data using the mean and std, and it assumes that the data follow a normal distribution. If the data deviate significantly from normality, the reliability of the Z-score results may be compromised. Experiments conducted by DeVries et al. in various study areas demonstrated that backscatter characteristics of natural areas during the same season can generally be considered normally distributed, and their water body monitoring based on the Z-score method achieved an overall accuracy of 90% (DeVries et al., 2020). To verify whether the backscatter intensity of building areas in time-series SAR images follows a normal distribution, this section employs kurtosis and skewness as statistical indicators.

Kurtosis is a statistical measure used to describe the shape of a data distribution. It indicates the degree to which a dataset contains outliers and tail risk. A normal distribution has a kurtosis value of 3. Subtracting 3 from the kurtosis yields excess kurtosis, which equals 0 for a perfectly normal distribution. The formula for calculating excess kurtosis is (DeCarlo, 1997):

$$\text{Kurtosis} = \frac{n(n+1) \sum (x_i - \mu)^4}{(n-1)(n-2)(n-3)\sigma^4} - \frac{3(n-1)^2}{(n-2)(n-3)} - 3 \quad (1)$$

where x_i represents the elements in the array and n is the sample size, while μ and σ denote the mean and standard deviation of the array, respectively. The sign of the excess kurtosis indicates whether the shape of the distribution is sharper or flatter compared to a normal distribution. A positive excess kurtosis suggests a distribution with heavier tails and a higher likelihood of outliers, whereas a negative excess kurtosis indicates fewer outliers and a flatter distribution.

Skewness is a statistical measure that quantifies the degree of asymmetry in a data distribution. For example, a normal distribution has a skewness value of 0. In cases of asymmetric distributions, if the tail is longer on the left side, it is referred to as left-skewed (negative skewness); if the tail is longer on the right, it is right-skewed (positive skewness). Skewness is

calculated by summing the cubed deviations from the mean and normalizing by the standard deviation. The formula for skewness is (Groeneveld and Meeden, 1984) :

$$\text{Skewness} = \frac{1}{n} \times \frac{\sum_{i=1}^N (x_i - \mu)^3}{\sigma^3} \quad (2)$$

180 where x_i represents the elements in the array and n is the sample size, while μ and σ denote the mean and standard deviation of the array, respectively. Skewness can be either positive or negative: right-skewed data has positive skewness, while left-skewed data have negative skewness. Skewness can be used to assess whether a dataset follows a normal distribution or another specific distribution, thereby providing a basis for statistical analysis or modeling.

To quantitatively assess the normality assumption underlying the Z-score transformation, we calculated the median
185 skewness and median excess kurtosis of the same-orbit Sentinel-1 backscatter time series for representative natural and urban land-cover classes during the non-flood reference period (Table 1). The selected classes included cropland, bare soil, open water, low-rise buildings, and high-density buildings. The diagnostic results indicate that the temporal backscatter distributions were not perfectly Gaussian. For the individual VV and VH backscatter channels, the median skewness ranged from 0.734 to 1.740, and the median excess kurtosis ranged from 0.048 to 2.596. All classes showed positive skewness, indicating that the
190 non-flood backscatter time series generally had a right-skewed distribution. The deviation from normality was relatively small for urban VH backscatter, especially in low-rise buildings, where the median skewness and excess kurtosis were 0.734 and 0.048, respectively, and in high-density buildings, where they were 0.953 and 0.560, respectively. This suggests that VH backscatter in built-up areas was comparatively stable during the reference period. In contrast, cropland showed larger deviations, particularly in VH polarization, with a median skewness of 1.740 and a median excess kurtosis of 2.596. This is
195 likely related to crop phenology, vegetation structure changes, and rainfall-induced surface moisture variations. The VV/VH ratio showed stronger departures from normality than the individual VV and VH channels, with median skewness ranging from 1.359 to 2.921 and median excess kurtosis ranging from 1.443 to 3.330. The largest skewness was observed for open water in VV/VH, indicating that ratio-based polarimetric variables can be more sensitive to low backscatter values and surface roughness variations. Therefore, the Z-score confidence interpretation in this study is applied primarily to the independently
200 standardized VV and VH backscatter channels rather than to the VV/VH ratio. These results show that the normality assumption is only approximately satisfied. Accordingly, the Z-score thresholds should not be interpreted as exact pixel-level flood probabilities in all land-cover types. Instead, they are used as standardized anomaly levels that retain statistical interpretability under approximate normality. This interpretation is particularly important for vegetation-covered areas and mixed urban pixels, where skewness and excess kurtosis indicate stronger deviations from a Gaussian distribution.

205

Table 1: Normality diagnostics of same-orbit Sentinel-1 backscatter during the non-flood reference period

Land-cover class	Polarization	Median skewness	Median excess kurtosis	Interpretation
Cropland	VV	1.049	0.564	Phenology may cause moderate deviations.
	VH	1.740	2.596	
	VV/VH	1.859	3.330	
Bare soil	VV	1.141	1.058	Expected to be relatively stable except after rainfall or tillage.
	VH	1.273	1.068	
	VV/VH	1.734	2.798	
Open water	VV	1.082	0.818	Wind-induced roughness may increase tails.
	VH	1.025	0.522	
	VV/VH	2.921	1.797	
Low-rise buildings	VV	1.109	0.806	Mixed pixels and local drainage may affect normality.
	VH	0.734	0.048	
	VV/VH	1.607	2.434	
High-density buildings	VV	1.109	0.929	Usually stable but sensitive to rainfall and double-bounce changes.
	VH	0.953	0.560	
	VV/VH	1.359	1.443	

Regarding the effect of sample size on skewness in normally distributed data, Doane et al. provided skewness coefficient thresholds corresponding to a 90% confidence level. Specifically, when the sample size is 40, the skewness should fall within the range of -0.594 to 0.594 (Doane, 2011). The Z-score values—accounting for differences in polarization mode, orbit parameters, and imaging mode—were then calculated from the mean and standard deviation of the time-series SAR images using the following formula:

$$Z_{pq,o,m}^{flood} = \frac{\sigma_{flood}^{pq,o,m} - \text{mean}(\sigma^{pq,o,m})}{\text{std}(\sigma^{pq,o,m})} \quad (3)$$

where Z represents the Z-score value, pq denotes the polarization mode (VV or VH), o indicates the orbit number, and m refers to the imaging mode. The normalized backscatter coefficient is standardized using the mean and standard deviation of the corresponding same-orbit reference time series. Under approximate normality, the Z-score can be associated with a nominal confidence level indicating the statistical extremeness of the flood-period observation relative to the reference distribution. For example, when |Z| exceeds 1.96, the observation lies outside the central 95% interval of the reference distribution. However, because Sentinel-1 backscatter may deviate from normality in phenologically dynamic vegetation and mixed urban pixels, the Z-score is interpreted in this study primarily as a standardized anomaly magnitude rather than an exact pixel-level flood probability. The expected anomaly direction differs by surface type: open-water flooding in natural areas usually produces negative anomalies, whereas urban inundation can produce positive anomalies through enhanced double-bounce scattering.

3.2 Flood monitoring classification tree

Using single-polarization SAR imagery alone yields limited accuracy in flood monitoring. Therefore, this study incorporates both VV and VH polarizations to provide multi-dimensional observational information and enhance the reliability of the flood mapping results. For example, during periods of heavy rainfall, rain and strong winds can increase water surface roughness, potentially leading to Bragg scattering dominance. This can strengthen the backscatter in co-polarized channels and obscure the presence of water bodies. In such cases, cross-polarization is less affected by surface roughness, offering additional evidence for flood delineation (Clement et al., 2018). In natural areas, the reference-baseline images used in this study were acquired 2–3 months prior to the flood event. If certain regions were continuously covered by water during this period, anomaly detection methods would struggle to identify them as flooded. To address this, auxiliary permanent surface water data from the GSW v1.4 dataset was incorporated to ensure the accuracy of flood extent delineation (Pekel et al., 2016). This section adopts the conceptual framework of decision tree classification, and by simultaneously considering dual-polarization information and surface water data, proposes a novel flood monitoring classification tree.

In this study, Equation (3) was applied pixel by pixel to both VV and VH images to generate their respective Z-score maps. Because VV and VH polarizations have different physical sensitivities to surface roughness, vegetation structure, and urban double-bounce scattering, the use of identical Z-score thresholds for the two polarization channels requires further clarification. In this study, the same nominal threshold was applied to VV and VH only after each polarization channel had been standardized independently using its own same-orbit temporal mean and standard deviation. Therefore, a given threshold value represents the same standardized anomaly magnitude within each polarization, rather than the same absolute backscatter change in dB. This treatment reduces the direct influence of the different dynamic ranges of VV and VH and allows the two polarizations to be combined within a statistically consistent anomaly-detection framework. For natural areas, floodwater generally causes a negative backscatter anomaly due to specular reflection; therefore, pixels with the VV and VH Z-score values lower than the negative threshold were identified as flood candidates. For building areas, urban inundation can enhance double-bounce scattering between water surfaces and vertical structures; therefore, pixels with the VV and VH Z-score values higher than the positive threshold were identified as urban flood candidates.

The classification tree algorithm begins by determining the land cover type of each pixel. In natural areas, the first step is to identify whether a pixel represents a seasonal water body based on its water occurrence frequency. Typically, a pixel is classified as a seasonal water body if its occurrence value exceeds 25, meaning the pixel is identified as water in more than 25% of the annual satellite observations. Such pixels are labeled in blue and excluded from further classification. For non-seasonal water bodies, if both the VV and VH Z-score values of a given pixel fall below a predefined threshold (th_n), the pixel is assigned a dual-polarization flood label. If only one of the polarizations falls below the threshold (th_n), the pixel is assigned a single-polarization flood label. In urban (building) areas, flood labels are similarly assigned based on whether the Z-score values of VV and VH are below a separate threshold (th_s). The same rule-based logic is applied to determine single- or dual-

polarization flood classifications. It is important to note that in building areas, the color coding for flood labels differs from that used in natural areas.

The same numerical thresholds are used for the independently standardized VV and VH Z-score maps because each polarization is normalized by its own temporal mean and standard deviation. Thus, an identical Z-score threshold represents a comparable standardized anomaly level within each polarization. Nevertheless, VV and VH have different sensitivities to wind-roughened water, vegetation structure, and urban double-bounce scattering. The GSW occurrence threshold of 25% was selected to remove recurrent seasonal/permanent water while retaining rarely inundated floodplain pixels. To evaluate the sensitivity of this choice, occurrence thresholds of 10%, 25%, 50%, and 75% were tested. Before classification, the 30 m GSW layer was co-registered to the Sentinel-1 processing grid. Because resolution mismatch and geolocation errors mainly affect narrow river margins, an additional uncertainty assessment excluded a one-pixel buffer around persistent-water boundaries during accuracy evaluation.

In SAR-based flood monitoring, even when filtering algorithms are applied during the preprocessing stage, speckle noise cannot be completely eliminated. As a result, the final flood mapping output often contains scattered misclassified pixels within large, homogeneous land cover areas. These fragmented misclassifications negatively impact the overall accuracy and may reduce the reliability of the flood monitoring results. To address this issue, this study adopts a method known as pixel aggregation. This technique specifically targets small misclassified regions. In particular, if a flood-classified area consists of four or fewer connected pixels, these pixels are reclassified as non-flood. The goal is to minimize the impact of noise-induced misclassification and improve the overall classification accuracy. This step serves as a post-processing refinement of the threshold-based classification results. Other commonly used post-processing methods include morphological operations such as erosion and dilation based on opening and closing. Through such post-processing techniques, small fragmented errors can be effectively removed, thereby enhancing the accuracy of flood monitoring outcomes (Borrelli et al., 2020). To avoid over-smoothing, this connected-pixel threshold was evaluated using 0, 2, 4, and 8 pixels. The ≤ 4 -pixel rule was retained as the operational setting because it suppresses isolated speckle-induced false alarms while preserving the main flood pattern; however, it may remove very small genuine puddles or narrow drainage features.

3.3 Accuracy verification of flood extraction

In this study, Sentinel-2 optical images and high-spatial-resolution unmanned aerial vehicle (UAV) optical images were used as validation data sources for flood extraction results in natural areas and urban (building) areas, respectively. The time difference between the Sentinel-2 images and the corresponding Sentinel-1 acquisition was 8–9 hours, while the UAV images had a 2-hour time difference. For water body extraction, the modified normalized difference water index (MNDWI) utilizes the green and short-wave infrared (SWIR) bands, which enhances the contrast between building areas and water bodies. This reduces confusion between the two and improves classification accuracy. The expression is as follows (Xu, 2006):

$$\text{MNDWI} = \frac{\text{Green} - \text{SWIR}}{\text{Green} + \text{SWIR}} \quad (4)$$

where, in the Sentinel-2 images, the green and SWIR bands correspond to bands 3 and 11, respectively. Since MNDWI is a normalized index, it also performs well in detecting water bodies located in shadowed areas.

In this study, high-confidence and medium-confidence flood pixels are both classified as flood pixels during accuracy assessment. Therefore, binary classification metrics are used for evaluation. The confusion matrix for binary classification is a 2×2 matrix, with the four elements being: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Common evaluation metrics based on these four elements include overall accuracy (OA), user accuracy (UA), and producer Accuracy (PA). In flood monitoring, the flood class requires special attention, and prediction accuracy may not always be the primary concern—particularly in cases where class distribution is highly imbalanced. The critical success index (CSI), which considers both PA and UA, is a more suitable metric for flood monitoring as it effectively reflects both the accuracy and stability of predictions. Therefore, this study adopts four evaluation metrics, including CSI. Their formulas are as follows:

To quantify uncertainty caused by the 8-9 h time difference between Sentinel-1 and Sentinel-2 acquisitions, the two available Sentinel-2-derived water masks before and after the SAR acquisition were used to define a temporal uncertainty envelope. Pixels that changed between the two optical masks were treated as flood-transition pixels. Accuracy metrics were then recalculated after excluding these dynamic pixels, allowing the effect of temporal mismatch on CSI and OA to be estimated.

For urban validation, UAV imagery was used as a high-confidence reference because of its high spatial resolution and approximately 2 h temporal difference from the Sentinel-1 acquisition. However, the UAV images cover only part of the urban domain. Therefore, the UAV-based accuracy assessment should be interpreted as a local high-resolution validation rather than a complete city-wide assessment.

$$PA = \frac{TP}{TP + FN} \quad (5)$$

$$UA = \frac{TP}{TP + FP} \quad (6)$$

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \quad (7)$$

$$CSI = \frac{TP}{TP + FP + FN} \quad (8)$$

4 Results and discussions

310 4.1 Scattering characteristics of natural areas and building areas in time-series SAR images

4.1.1 Natural areas

In this study, natural areas refer to regions with land cover types such as water bodies, bare soil, cropland, and forests—i.e., non-artificial surfaces. These areas exhibit different dominant scattering mechanisms in SAR images. For water-covered areas, the dominant scattering mechanism is specular reflection. However, ripples or waves caused by wind or other disturbances can alter the surface roughness of the water, introducing components of surface scattering or Bragg scattering. Although water bodies generally exhibit low backscatter values across all four polarization modes, different polarizations vary in their sensitivity to surface roughness. For example, in dual-polarization Sentinel-1 images, roughened water surfaces tend to show a greater increase in backscatter under VV polarization than under VH polarization. As shown in Fig. 2, some pixels in the central river area reach backscatter intensities of -13 dB, which are close to those of surrounding non-water areas. This can lead to a reduced extent of the identified water body.

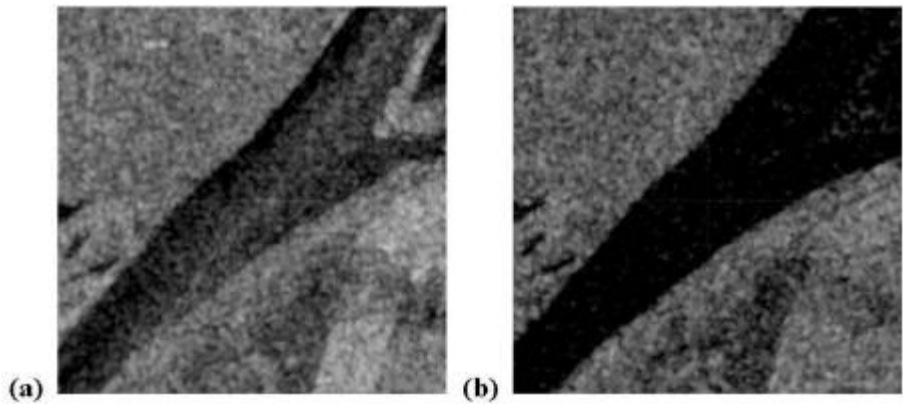


Figure 2: Comparison of different polarization images on the surface of roughened water bodies. (a) shows VV polarization image. (b) shows VH polarization image.

The backscatter from bare soil areas contains rich soil information, including soil moisture and soil texture. In arid desert regions, low soil moisture results in low internal charge polarization and a relatively small dielectric constant, allowing microwaves to easily penetrate the surface soil. In such cases, volume scattering is the dominant mechanism. As soil moisture increases, the dielectric constant also increases, and surface scattering becomes the dominant type. Different soil textures can also influence the dominant scattering mechanism—for example, fine-grained sand is more prone to volume scattering than blocky clay. Croplands are key areas of interest in flood monitoring. Compared to water bodies and bare soils, the backscatter coefficient in croplands shows distinct seasonal variations corresponding to crop growth stages.

Crop growth is generally divided into four stages: germination, vegetative, reproductive, and maturity. During the germination stage, crops are sparsely distributed, and the surface remains largely bare soil, resulting in backscatter characteristics similar to those of bare land. In the vegetative stage, crops grow rapidly. Increased stem height and higher leaf

number and density lead to a sharp rise in the number of scattering elements, significantly enhancing volume scattering—
335 particularly in VH polarization, where the backscatter coefficient increases notably. For VV polarization, the increase is less
pronounced because crop stems are oriented nearly perpendicular to the ground, leading to a weaker response. During the
reproductive and maturity stages, crop changes primarily involve fruit development, with little change in height or leaf number.
The newly formed fruits serve as additional scattering elements, enhancing volume scattering. As a result, backscatter intensity
continues to rise but remains relatively stable overall. By comparing three SAR images taken during different growth stages
340 and the flood period, it is evident that there is a significant difference in backscatter intensity between the vegetative and
maturity stages. As crops mature, backscatter intensity increases markedly. However, the intensity remains relatively stable
between the vegetative stage and the flood period, making it more challenging to distinguish flooding from normal growth
during that time.

4.1.2 Building areas

345 The scattering characteristics of urban (building) areas are highly complex and influenced by both imaging parameters (such
as polarization mode, incidence angle, and resolution) and surface attributes (including building orientation, layout, and surface
roughness). Densely arranged man-made structures can give rise to single, double, triple, or even multiple scattering events.
As the number of scattering interactions increases, the signal's round-trip path length becomes longer, resulting in severe signal
mixing among adjacent buildings. This poses significant challenges for interpreting SAR images in urban environments. Due
350 to the large distance between spaceborne radar and ground targets, the incident radar beams can be approximated as parallel
rays, simplifying certain aspects of geometric analysis but not mitigating the inherent complexity of urban backscatter.

The angle between building walls and the satellite's orbital direction can significantly influence the strength of double-
bounce scattering. When the orbital direction is parallel to the building walls—meaning the radar beam plane is perpendicular
to the wall surface—the double-bounce scattering is strongest. In this case, the echo generated by specular reflection lies in
355 the same plane as the incident wave. When there is an angular offset between the orbital direction and the wall, the radar cross-
section (RCS) energy is primarily composed of surface scattering, and the backscatter intensity decreases rapidly as the angle
increases. In change detection or time-series analysis, it is essential to ensure that all images are acquired with the same orbit
direction (either ascending or descending) to avoid significant changes in backscatter caused by orbital differences, which
could reduce detection or analysis accuracy. During urban flooding, the land surrounding buildings becomes water-covered,
360 leading to specular reflection as the dominant scattering type and a reduction in backscatter intensity. Assuming the building
wall roughness remains unchanged, radar echoes are strengthened by energy reflected from the water surface onto the walls,
resulting in a significant increase in backscatter intensity. Since the contributions from the sidewalls and rooftops remain
largely unchanged, the transition zones between water and buildings may show a slight decrease in intensity, as seen in Fig. 3.

When other man-made structures are located near buildings, multiple scattering between sidewalls becomes complex and
365 difficult to predict or model. Due to the increased round-trip path length, these effects are typically observed on the right side
of the “bright line” region in the SAR image. Therefore, urban flooding tends to enhance backscatter intensity in building areas,

but the magnitude of this enhancement is highly uncertain. Fig. 3 also illustrates changes in urban backscatter intensity following rainfall that occurred between July 19 and 22, 2021, with 24-hour precipitation totals of 103 mm, 380 mm, 398 mm, and 76 mm, respectively. The false-colour composite image assigns the VV-polarized SAR images acquired on July 27, July 3, and June 21 to the RGB channels, respectively. In the image, blue-green areas represent open water bodies in natural regions, as the values in the green and blue channels are significantly higher than those in the red channel. This suggests that the dominant scattering has shifted from volume scattering (typical of mature wheat fields) to specular reflection, indicating the crops were inundated by floodwaters.

Bright-coloured areas with regular texture patterns indicate building regions. A comparison of the building areas northwest and southeast of the river reveals the significant impact of flooding on backscatter intensity. The southeastern urban area appears grey-white in the image, indicating that the backscatter intensity across the three acquisition dates is relatively consistent, with no significant change. In contrast, the northwestern urban area appears red, indicating a sharp increase in brightness in the red channel—i.e., a significant rise in backscatter intensity during the flood period. This suggests severe urban flooding, which is further supported by the presence of extensive surrounding surface water.

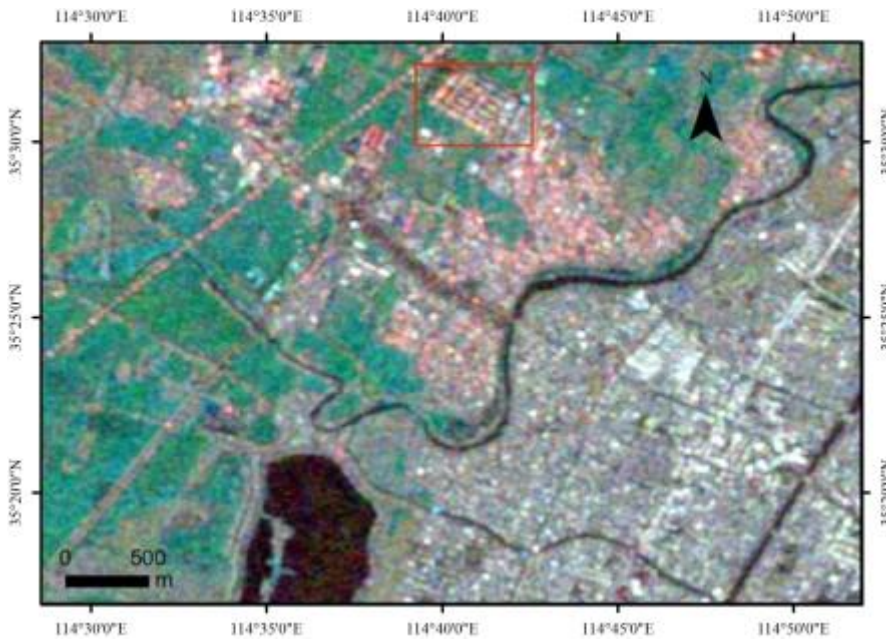


Figure 3: The influence of heavy rainfall on backscatter intensity in building areas. Sentinel-1 SAR composite processed by the authors (Sentinel-1 data courtesy of ESA Copernicus).

A close examination of the building area in the northwest reveals significant variations in building structures, layout patterns, and densities. For instance, the area outlined in red consists of uniformly arranged high-rise buildings, while most of the surrounding areas are densely packed with low-rise structures. Despite this variation, the region still appears predominantly red in the image, indicating that when standing water is present around buildings, the backscatter intensity increases significantly, regardless of whether the buildings are low-rise, mid-rise, or high-rise.

According to GPM rainfall data, this trend becomes even more apparent in time-series SAR images, as shown in Fig. 4. In the absence of flooding, the VV-polarized backscatter intensity in building areas ranges from -8.6 dB to -4.0 dB, while the VH-polarized intensity ranges from -13.7 dB to -9.8 dB, both exhibiting relatively stable fluctuations. This stability suggests that, compared to natural areas, building areas are less affected by seasonal variation, allowing for greater flexibility in time-series analyses without the need to constrain the images to the same season. This is particularly advantageous for studying changes and development trends in building areas. Once flooding occurs, both VV and VH backscatter intensities increase sharply—by more than 6 dB—further indicating that changes in the roughness of surrounding surfaces disrupt the previously consistent backscatter signal of building areas. Therefore, using time-series SAR images to monitor urban flooding is both feasible and effective.

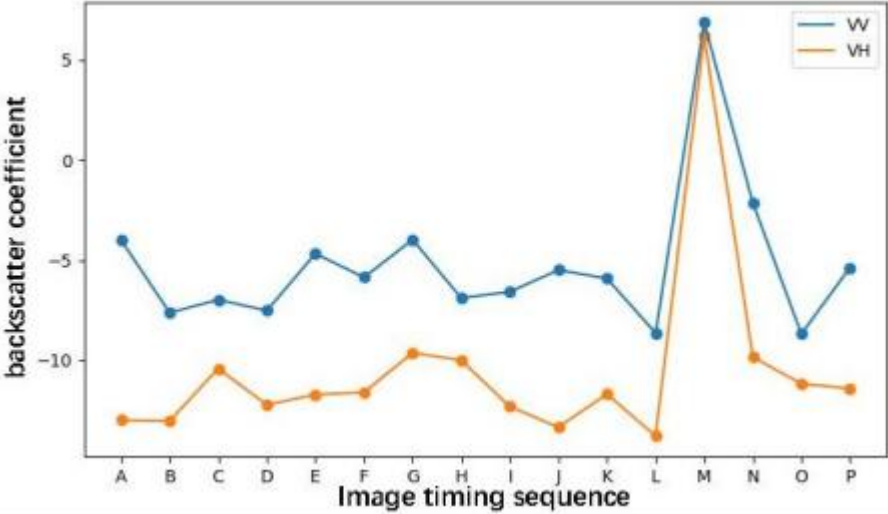
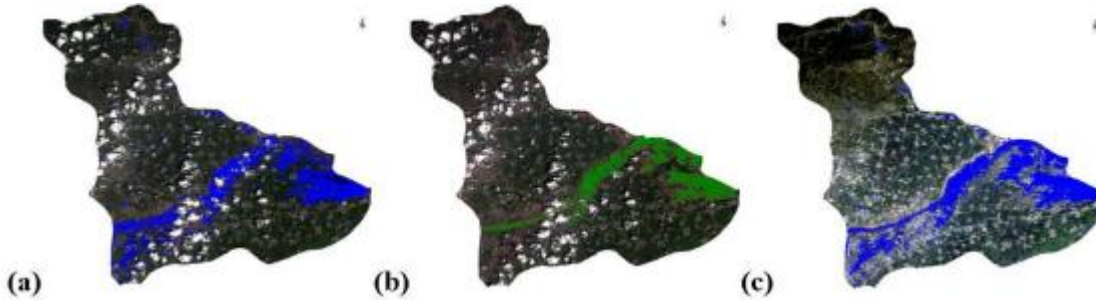


Figure 4: Trend of temporal SAR images (VV and VH polarization) in building areas within the red rectangular frame in Fig. 3.

4.2 Flood monitoring results and analysis in natural areas

When using the Z-score method for flood monitoring, it is necessary to set a threshold to distinguish between anomalous (flooded) and non-anomalous (non-flooded) pixels. However, unlike traditional thresholding methods, thresholds based on Z-scores have a clear statistical confidence level. To compare the accuracy of flood monitoring results under different confidence levels, this section defines three Z-score thresholds: -2, -2.5, and -3. According to the standard normal distribution table, these thresholds correspond to confidence levels of 95.4%, 98.8%, and 99.7%, respectively. The Z-score images generated using reference-baseline images and the time-series SAR flood monitoring algorithm were clipped to match the coverage area of the validation images. The Z-score values were then classified using three thresholds, as shown in Fig. 5. It is important to note that seasonal water bodies were excluded during the accuracy assessment. This is because such water bodies consistently exhibit low backscatter intensity in time-series SAR images, whereas time-series monitoring primarily aims to highlight newly emerged surface water.

410 In Fig. 5, different colours are used to represent the flood monitoring results under different Z-score thresholds. As shown, lowering the confidence level introduces more speckled flood pixels but also captures more water bodies. For instance, noticeable pink pixels appear in the northern and eastern parts of Weihui under less strict thresholds, whereas using a threshold of -3 results in more missed detections in these areas. A comparison between Fig. 5a and 5b reveals a significant difference between the flood extent identified by Sentinel-2 on the 26th and that detected by SAR on the 27th, particularly in the floodplains upstream and downstream of the Wei River. These differences are also reflected in the accuracy assessment, where the CSI and PA show discrepancies of up to 17% and 23%, respectively. The OA shows a slight increase, primarily because the flood and non-flood classes in Weihui are unevenly distributed, resulting in a high proportion of TN in the confusion matrix.



420 **Figure 5:** Comparison of monitoring results in the flood area of Weihui City based on the Z-score method. (a), 07/26/2021, (b), 07/27/2021, (c), 07/31/2021. The blue or green areas in the picture represent the areas affected by the flood.

A comparison between the optical water extraction results from July 26 and July 31 reveals significant changes in the flood extent within Weihui City during this period. The flood areas in the upstream and midstream regions contracted, while the flooded areas in the downstream region expanded. As shown clearly in Fig. 6, the monitoring results from July 27 are highly consistent with the ground truth from July 31 in the downstream floodplain, whereas there is a notable difference when compared to the July 26 ground truth. This highlights the dynamic nature of flood disasters. When evaluating the accuracy of flood monitoring algorithms, it is crucial to consider both timeliness—the validation images should be as close in time as possible to the SAR acquisition—and sample size, using as much validation data as possible to ensure accurate and reliable assessment of the algorithm's performance. In the accuracy evaluations using the two Sentinel-2 scenes as ground truth, the proposed SAR-based flood monitoring algorithm achieved over 92% overall accuracy in both cases and showed strong consistency with the water extent derived from both the July 26 and July 31 ground truths. Overall, while lowering the confidence level introduced more speckled flood pixels, which led to a noticeable decline in UA, the CSI—which places greater emphasis on the correct identification of flood areas—increased to varying degrees. Therefore, a Z-score threshold of -2 can be considered appropriate for flood detection in natural areas. At the 95% confidence level, both study areas achieved approximately 90% OA and CSI values above 60%. In conclusion, the method proposed in this study demonstrated high accuracy within the study areas, indicating strong reliability and stability for flood monitoring applications.



Figure 6: Comparison of monitoring results in the downstream flood monitoring area of Weihui City based on the Sentinel-2 images (Sentinel-2 data courtesy of ESA Copernicus). (a), 07/26/2021, (b), 07/27/2021, (c), 07/31/2021. The blue or green areas in the picture represent the areas affected by the flood.

4.3 Flood monitoring results and analysis in building areas

In urban flood monitoring, it is also necessary to determine thresholds for identifying anomalies. This study found that the frequency distribution curves of Z-score images in building areas do not exhibit a bimodal pattern, indicating that the Otsu method (maximum inter-class variance) is not suitable for determining accurate thresholds in this context. Therefore, this section adopts the same confidence-level-based thresholds used for natural areas to monitor urban flooding, with threshold values set at 2, 2.5, and 3. In natural areas, using VV and VH polarizations from Sentinel-1 images helps mitigate interference from rough water surfaces, yielding better monitoring performance than single-polarization approaches. However, few studies have applied Sentinel-1 images for urban flood detection. Thus, this section first evaluates the performance of VV, VH, and dual-polarization combinations under the same threshold conditions in monitoring urban flooding.

The UAV images of Weihui City offer not only ultra-high spatial resolution and oblique viewing advantages, but also excellent timeliness. Given the 2-hour difference between the UAV and SAR acquisition times, it can reasonably be assumed that the flood extent remained largely unchanged during this interval. In this section, this validation dataset is used for a detailed analysis of the urban flood detection results based on time-series Sentinel-1 SAR images (Fig. 7). This study found that lower threshold values result in a significantly higher number of detected flood pixels compared to higher thresholds. In many areas, flood pixels identified at lower thresholds effectively filled the gaps and voids present in the higher-threshold results. Overall, the flood regions identified using lower thresholds exhibited better spatial continuity. Aside from a slight decline in UA, all other accuracy metrics showed improvements to varying degrees. Notably, the CSI—a key indicator of flood detection performance—achieved a 7.88% increase in accuracy.

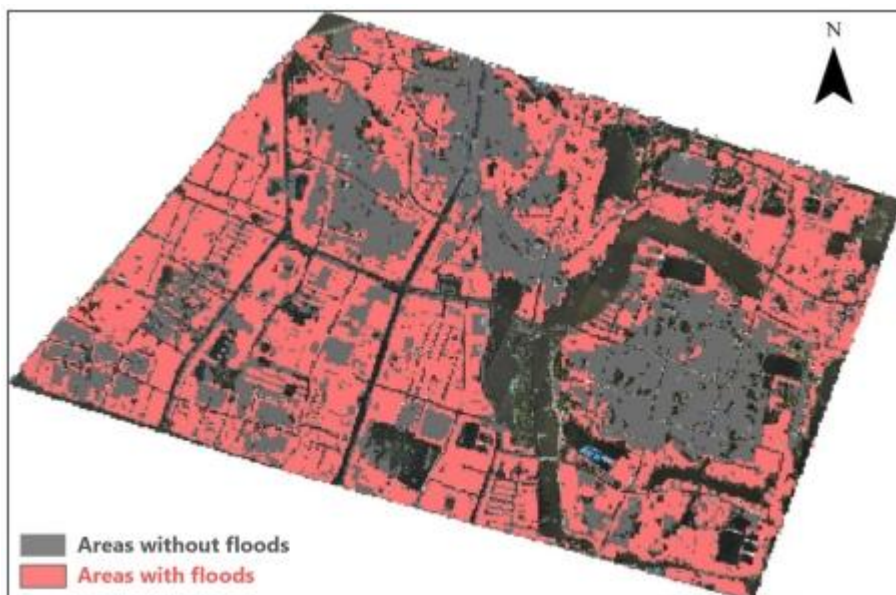


Figure 7: Flood monitoring results in the building area of Weihui City based on Sentinel-1 images and the Z-score method (Sentinel-1 data courtesy of ESA Copernicus). Grey indicates that the area has not been affected by the flood; red indicates that the area has been affected by the flood. The UAV imaging range is shown in Figure 1.

By comparing the identified flood regions with the reference data (Figure 8), we observed substantial discrepancies in several building areas. This section focuses on two representative regions where the algorithm failed to detect urban inundation effectively even under a relatively low threshold, indicating that the expected flood-induced backscatter enhancement was weak. Orthorectified images alone are insufficient to explain this phenomenon. Therefore, the UAV oblique views shown in Figures 8a and 8b were examined to better understand the local building and inundation conditions.

The UAV images show that these areas are low-lying and are dominated by single-story or low-rise residential buildings. In these locations, floodwater nearly reached the rooftops, substantially reducing the relative height difference between buildings and the surrounding water surface. Under such conditions, the radar return is dominated mainly by rooftop surface scattering, while wall–water double-bounce scattering becomes weak or absent. As a result, severe urban inundation may still be missed by intensity-based SAR detection methods.

This failure mechanism indicates that the building–water height difference is a key factor controlling the detectability of urban inundation in SAR imagery. When floodwater remains substantially lower than building height, wall–water double-bounce scattering can be enhanced, increasing backscatter intensity and facilitating flood detection. In contrast, when floodwater approaches the roof level of low-rise buildings, especially in dense residential areas with narrow building spacing, the side-looking SAR geometry may not produce a clear double-bounce response. This leads to weak flood-related backscatter anomalies and increases the risk of missed detection.

High-spatial-resolution DEM or digital surface model (DSM) data, building footprints, and building-height information would be useful for identifying such failure-prone areas by estimating local building–water height differences. However, the

480 Copernicus DEM used in this study has a spatial resolution of 30 m and cannot resolve individual buildings, narrow streets, or local elevation differences. Therefore, the proportion of building areas affected by this failure mechanism was not quantitatively estimated in the present study. This limitation is further discussed as a future research direction in Section 5.



Figure 8: UAV optical images when floods occurred in Weihui City. (a) represents areas severely inundated by floods; (b) represents areas with moderate flood inundation. UAV image source: authors' field survey.

4.4 Comparative analysis with other studies

To further evaluate the reliability of the proposed method, we compared it with four baseline methods, including VV-only Z-score detection, VH-only Z-score detection, simple image differencing, and Otsu thresholding (Table 2). All methods were evaluated using the same Sentinel-1 acquisitions and validation masks to ensure a fair comparison.

490 In natural areas, Otsu thresholding achieved the highest CSI and OA, with values of 63% and 92%, respectively. This result indicates that histogram-based thresholding can perform well in open floodplain environments where flooded and non-flooded pixels show relatively clear backscatter separation. The proposed dual-polarization Z-score tree achieved slightly lower but comparable accuracy in natural areas, with a CSI of 60% and an OA of 90%. Compared with VV-only Z-score detection, the proposed method improved CSI and OA by 9 and 10 percentage points, respectively. Compared with VH-only Z-score
495 detection, the improvements were 4 percentage points for both CSI and OA. Compared with simple image differencing, the proposed method improved CSI by 4 percentage points and OA by 5 percentage points. These results indicate that the proposed dual-polarization framework retains competitive performance in natural floodplain areas while reducing the limitations of single-polarization and simple change-detection methods.

The advantage of the proposed method is more evident in building areas. The proposed method achieved the highest CSI and OA, reaching 62% and 73%, respectively. In comparison, VV-only and VH-only Z-score detection achieved CSI/OA values of 52%/63% and 54%/65%, respectively, indicating that single-polarization information is insufficient to fully capture the complex scattering response of urban inundation. Simple image differencing performed worse in building areas, with a CSI of 45% and an OA of 56%, because non-flood-related backscatter changes caused by rainfall, surface wetness, and urban structural heterogeneity can be incorrectly interpreted as flood-induced changes. Otsu thresholding showed the largest

505 performance degradation, with CSI and OA decreasing to 39% and 50%, respectively. This confirms that Otsu thresholding is unstable in building areas where the Z-score distribution does not exhibit a clear bimodal pattern.

The Sentinel-1 Global Flood Monitoring product is an operational near-real-time flood mapping service of the Copernicus Emergency Management Service. It processes incoming Sentinel-1 SAR acquisitions globally and derives flood extent using an ensemble of three independent automated flood-mapping algorithms. The final observed flood extent is generated by a majority-decision ensemble and is accompanied by reference water, exclusion mask, likelihood, and advisory-flag layers. Although GFM provides global near-real-time flood information and a historical Sentinel-1 archive from 2015 onward, its effective availability and reliability depend on the Sentinel-1 acquisition schedule and on exclusion-mask conditions such as topographic distortion, radar shadow, dense vegetation, building areas, and low-sensitivity surfaces. Therefore, for urban flood mapping in Weihui City, the Sentinel-1 Global Flood Monitoring product appears to be ineffective. The accuracy of flood extraction from the natural areas of this product is slightly lower than that of the method used in this study.

Overall, the proposed method provided the most balanced performance across both natural and urban environments. Its mean CSI and OA across the two area types were 61% and 81.5%, respectively, which were higher than those of all baseline methods. More importantly, the proposed method had the highest worst-case performance, with a minimum CSI of 60% and a minimum OA of 73% across the two environments. In contrast, although Otsu thresholding performed best in natural areas, its CSI and OA decreased sharply in building areas. These results demonstrate that the proposed dual-polarization Z-score classification tree is not only accurate in open floodplain areas but also more reliable and transferable for complex urban inundation mapping.

Table 2: Accuracy comparison among different methods

Method	Natural area (CSI/OA)	Building area (CSI/OA)	Data availability
VV-only Z-score	51%/80%	52%/63%	Same Sentinel-1 acquisitions and validation masks
VH-only Z-score	56%/86%	54%/65%	
Simple image differencing	56%/85%	45%/56%	Reference and flood-period SAR images
Otsu thresholding	63%/92%	39%/50%	Expected to be unstable if the histogram is not clearly bimodal
Sentinel-1 Global Flood Monitoring product	57%/85%	—/—	Most of the area of the city has been masked
Proposed dual-polarization Z-score tree	60%/90%	62%/73%	This study

5 Deficiencies and prospects

525 Time-series SAR images offer significant advantages in flood monitoring, and the method proposed in this study achieved high detection accuracy in both natural areas and building areas. However, the potential of time-series SAR data has not yet been fully realized, and several aspects warrant further research and refinement:

(1) Although forests and building areas exhibit similar trends in backscatter intensity during flood events, this study did not address under-canopy flood detection due to a lack of validation data. Future research can explore this topic in depth, analysing
530 how factors such as phenological stages and precipitation events affect the consistency of backscatter intensity in forested regions. Based on this, a more robust under-canopy flood monitoring algorithm could be developed. Additionally, the proposed method holds value for flood assessment in agricultural areas; for instance, in croplands with taller vegetation, more accurate estimates of flooded area may be achievable.

(2) In urban flood detection, the accuracy of the proposed method is strongly influenced by building density, building height,
535 street width, and the relative height difference between buildings and surrounding floodwater. When buildings are densely packed or when the floodwater level approaches the height of low-rise buildings, wall–water double-bounce scattering becomes weak or absent, and the radar signal may be dominated by rooftop surface scattering. This can lead to missed detection in severely inundated low-rise residential areas. The Copernicus DEM used in this study has a spatial resolution of 30 m, which is useful for describing general low-lying terrain but insufficient for resolving individual buildings, narrow streets, and
540 building–water height differences. In future work, high-spatial-resolution DEM or DSM data, building footprints, and building-height information will be introduced to estimate local building–water height differences, identify areas where double-bounce scattering is likely to be suppressed, and further evaluate the accuracy and limitations of SAR-based urban inundation extraction.

(3) High-frequency flood monitoring can help generate more precise maps of flood-affected areas and provide timely support
545 for disaster response. Since the proposed method allows for sensor-specific or configuration-specific adjustments by independently computing baseline statistics, it is feasible to incorporate multi-source time-series SAR images to increase the observation frequency during flood events. Furthermore, SAR images acquired at different wavelengths may offer additional insights for flood monitoring—for example, L-band SAR from satellites such as LuTan-1 may offer greater advantages in detecting under-canopy flooding.

550 6 Conclusions

Based on an analysis of the C-band SAR backscatter characteristics across different land cover types, this study proposes a novel method that integrates various auxiliary datasets to enable both large-scale flood monitoring and urban inundation detection. The main conclusions are as follows: Analysis of the scattering characteristics of time-series SAR images showed that in natural areas, vegetation growth status significantly affects intensity consistency, whereas in building areas, rainfall-
555 induced enhancement of double-bounce scattering can also disrupt intensity stability. Through the analysis of time-series

backscatter intensity in urban regions, a Z-score-based flood classification tree was developed. The method calculates Z-score images using both reference and flood-period SAR images, enabling standardized anomaly levels with statistical interpretability under approximate normality and assigning operationally transparent meanings to flood-detection thresholds. The combination of VV and VH polarizations within the classification tree further enhances the reliability of flood identification. In the case of flood monitoring in Weihui City, the method demonstrated strong performance: the CSI and OA in natural areas reached 60% and 90%, respectively, while in building areas, the CSI and OA reached 62% and 73%, respectively. The proposed method shows clear advantages in accurately classifying flood-affected areas and possesses the capability to simultaneously monitor widespread flooding and urban inundation. Therefore, it holds great potential for application in disaster prevention, mitigation, and emergency response.

565

Code and data availability.

The database supporting this study is not publicly accessible because it contains sensitive, site-specific ecological information. Unrestricted disclosure of these data could expose sensitive ecological locations and increase the risk of disturbance, inappropriate use, or other adverse impacts on the associated ecological resources. The data may be made available by the corresponding author to qualified researchers for legitimate academic purposes upon reasonable request. Requests will be assessed based on the intended use of the data, and access will be subject to appropriate conditions to prevent the disclosure or misuse of sensitive ecological information. The code used for the statistical analyses may also be obtained from the corresponding author upon reasonable request. Before release, the code will be checked to remove any sensitive site identifiers or other information linked to the restricted database.

575

Author contributions.

JZ: Project lead, Study design, Data curation, Writing, Reviewing. RX: Formal analysis, Writing, Reviewing. FL: Data curation, Formal analysis, Reviewing.

580 ***Competing interests.***

The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note:

Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

585

References

- Alvarez-Vanhard, E., Corpetti, T., and Houet, T.: UAV and satellite synergies for optical remote sensing applications: a literature review, *Sci. Remote Sensing*, 3, 100019, <https://doi.org/10.1016/j.srs.2021.100019>, 2021.
- Amitrano, D., Di Martino, G., Di Simone, A., and Imperatore, P.: Flood detection with SAR: a review of techniques and datasets, *Remote Sens.*, 16, 656, <https://doi.org/10.3390/rs16040656>, 2024.
- Anusha, N. and Bharathi, B.: Flood detection and flood mapping using multi-temporal synthetic aperture radar and optical data, *The Egyptian Journal of Remote Sensing and Space Science*, 23, 207–219, <https://doi.org/10.1016/j.ejrs.2019.01.001>, 2020.
- Barredo, J. I.: Major flood disasters in Europe: 1950–2005, *Nat. Hazards*, 42, 125–148, <https://doi.org/10.1007/s11069-006-9065-2>, 2007.
- Borrelli, P., Robinson, D. A., Panagos, P., Lugato, E., Yang, J. E., Alewell, C., Wuepper, D., Montanarella, L., and Ballabio, C.: Land use and climate change impacts on global soil erosion by water (2015–2070), *Proc. Natl. Acad. Sci. U. S. A.*, 117, 21994–22001, <https://doi.org/10.1073/pnas.2001403117>, 2020.
- Cheadle, C., Vawter, M. P., Freed, W. J., and Becker, K. G.: Analysis of microarray data using Z score transformation, *J. Mol. Diagn.*, 5, 73–81, [https://doi.org/10.1016/S1525-1578\(10\)60455-2](https://doi.org/10.1016/S1525-1578(10)60455-2), 2003.
- Cian, F., Marconcini, M., Ceccato, P., and Giupponi, C.: Flood depth estimation by means of high-resolution SAR images and lidar data, *Nat. Hazards Earth Syst. Sci.*, 18, 3063–3084, <https://doi.org/10.5194/nhess-2018-158>, 2018.
- Clement, M. A., Kilsby, C. G., and Moore, P.: Multi-temporal synthetic aperture radar flood mapping using change detection, *J. Flood Risk Manag.*, 11, 152–168, <https://doi.org/10.1111/jfr3.12303>, 2018.
- Cruz, H., Véstias, M., Monteiro, J., Neto, H., and Duarte, R. P.: A review of synthetic-aperture radar image formation algorithms and implementations: a computational perspective, *Remote Sens.*, 14, 1258, <https://doi.org/10.3390/rs14051258>, 2022.
- DeCarlo, L. T.: On the meaning and use of kurtosis, *Psychol. Methods*, 2, 292, <https://doi.org/10.1037/1082-989X.2.3.292>, 1997.
- DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J. W., and Lang, M. W.: Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine, *Remote Sens. Environ.*, 240, 111664, <https://doi.org/10.1016/j.rse.2020.111664>, 2020.
- Doane, D. P. and Seward, L. E.: Measuring skewness: a forgotten statistic?, *J. Stat. Educ.*, 19, <https://doi.org/10.1080/10691898.2011.11889611>, 2011.
- Du, W., Xia, Q., Cheng, B., Xu, L., Chen, Z., Zhang, X., Huang, M., and Chen, N.: Flood inundation probability estimation by integrating physical and social sensing data: case study of 2021 heavy rainfall in Henan, China, *Remote Sens.*, 16, 2734, <https://doi.org/10.3390/rs16152734>, 2024.

- 620 Fernandez, P., Mourato, S., Moreira, M., and Pereira, L.: A new approach for computing a flood vulnerability index using cluster analysis, *Physics and Chemistry of the Earth, Parts A/B/C*, 94, 47–55, <https://doi.org/10.1016/j.pce.2016.04.003>, 2016.
- Geudtner, D., Torres, R., Snoeij, P., Davidson, M., and Rommen, B.: Sentinel-1 system capabilities and applications, in: *Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium*, 1457–1460, 625 <https://doi.org/10.1109/IGARSS.2014.6946711>, 2014.
- Groeneveld, R. A. and Meeden, G.: Measuring skewness and kurtosis, *Journal of the Royal Statistical Society Series D: The Statistician*, 33, 391–399, 1984.
- Jongman, B., Ward, P. J., and Aerts, J. C.: Global exposure to river and coastal flooding: long-term trends and changes, *Global Environmental Change*, 22, 823–835, <https://doi.org/10.1016/j.gloenvcha.2012.07.004>, 2012.
- 630 Kang, W., Xiang, Y., Wang, F., Wan, L., and You, H.: Flood detection in Gaofen-3 SAR images via fully convolutional networks, *Sensors*, 18, 2915, <https://doi.org/10.3390/s18092915>, 2018.
- Kummerow, C., Simpson, J., Thiele, O., Barnes, W., Chang, A., Stocker, E., Adler, R., Hou, A., Kakar, R., and Wentz, F.: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit, *J. Appl. Meteorol.*, 39, 1965–1982, [https://doi.org/10.1175/1520-0450\(2001\)040%3C1965:TSOTTR%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040%3C1965:TSOTTR%3E2.0.CO;2), 2000.
- 635 Kunkel, K. E.: North American trends in extreme precipitation, *Nat. Hazards*, 29, 291–305, <https://doi.org/10.1023/A:1023694115864>, 2003.
- Li, H., Zhao, J., Yan, B., Yue, L., and Wang, L.: Global DEMs vary from one to another: an evaluation of newly released Copernicus, NASA and AW3D30 DEM on selected terrains of China using ICESat-2 altimetry data, *Int. J. Digit. Earth*, 15, 1149–1168, <https://doi.org/10.1080/17538947.2022.2094002>, 2022.
- 640 Liu, B., Li, Y., Zhang, Y., Wu, Z., Pan, Y., and Li, M.: Multi-criteria decision-making method for emergency shelter site selection considering flood risk: a case study of Zhuhai, China, *Cleaner Eng. Technol.*, 24, 100892, <https://doi.org/10.1016/j.clet.2025.100892>, 2025.
- Liu, X., Huang, Y., Xu, X., Li, X., Li, X., Ciais, P., Lin, P., Gong, K., Ziegler, A. D., and Chen, A.: High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015, *Nat. Sustain.*, 3, 564–570, [https://doi.org/10.1038/s41893-](https://doi.org/10.1038/s41893-020-0521-x) 645 [020-0521-x](https://doi.org/10.1038/s41893-020-0521-x), 2020.
- Martinis, S., Kuenzer, C., Wendleder, A., Huth, J., Twele, A., Roth, A., and Dech, S.: Comparing four operational SAR-based water and flood detection approaches, *Int. J. Remote Sens.*, 36, 3519–3543, <https://doi.org/10.1080/01431161.2015.1060647>, 2015.
- Martinis, S., Plank, S., and Ćwik, K.: The use of Sentinel-1 time-series data to improve flood monitoring in arid areas, *Remote 650 Sens.*, 10, 583, <https://doi.org/10.3390/rs10040583>, 2018.
- Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., and Papathanassiou, K. P.: A tutorial on synthetic aperture radar, *Ieee Geosci. Remote Sens. Mag.*, 1, 6–43, <https://doi.org/10.1109/MGRS.2013.2248301>, 2013.

- Nhangumbe, M., Nascetti, A., and Ban, Y.: Multi-temporal Sentinel-1 SAR and Sentinel-2 MSI data for flood mapping and damage assessment in Mozambique, *Isprs Int. J. Geo-Inf.*, 12, 53, <https://doi.org/10.3390/ijgi12020053>, 2023.
- 655 Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, *Nature*, 540, 418–422, <https://doi.org/10.1038/nature20584>, 2016.
- Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., and Hong, Y.: Have satellite precipitation products improved over the last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets, *Remote Sens. Environ.*, 240, 111697, <https://doi.org/10.1016/j.rse.2020.111697>, 2020.
- 660 Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., and Slayback, D. A.: Satellite imaging reveals increased proportion of population exposed to floods, *Nature*, 596, 80–86, <https://doi.org/10.1038/s41586-021-03695-w>, 2021.
- Tingsanchali, T.: Urban flood disaster management, *Procedia Engineering*, 32, 25–37, <https://doi.org/10.1016/j.proeng.2012.01.1233>, 2012.
- 665 Torres, R., Snoei, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., and Brown, M.: GMES Sentinel-1 mission, *Remote Sens. Environ.*, 120, 9–24, <https://doi.org/10.1016/j.rse.2011.05.028>, 2012.
- Tripathi, G., Pandey, A. C., Parida, B. R., and Kumar, A.: Flood inundation mapping and impact assessment using multi-temporal optical and SAR satellite data: a case study of the 2017 flood in Darbhanga district, Bihar, India, *Water Resour Manage*, 34, 1871–1892, <https://doi.org/10.1007/s11269-020-02534-3>, 2020.
- 670 Twele, A., Cao, W., Plank, S., and Martinis, S.: Sentinel-1-based flood mapping: a fully automated processing chain, *Int. J. Remote Sens.*, 37, 2990–3004, <https://doi.org/10.1080/01431161.2016.1192304>, 2016.
- Voigt, S., Martinis, S., Zwenzner, H., Hahmann, T., Twele, A., and Schneiderhan, T.: Extraction of flood masks using satellite-based very-high-resolution SAR data for flood management and modeling, in: *RIMAX Contributions at the 4th International Symposium on Flood Defence (ISFD4), Managing Flood Risk, Reliability & Vulnerability*, Toronto, Canada, 6–8 May 2008.
- 675 Voigt, S., Giulio-Tonolo, F., Lyons, J., Kučera, J., Jones, B., Schneiderhan, T., Platzeck, G., Kaku, K., Hazarika, M. K., and Czarán, L.: Global trends in satellite-based emergency mapping, *Science*, 353, 247–252, <https://doi.org/10.1126/science.aad8728>, 2016.
- Wang, T. and Tong, C.: An analytical model of double scattering for SAR imaging of urban structures, *Optik*, 171, 484–491, <https://doi.org/10.1016/j.ijleo.2018.06.084>, 2018.
- 680 Wang, Z., Qin, Y., Zhang, Q., Li, Y., Liu, J., Yuan, X., Zhao, L., Li, M., and Cao, S.: China’s GaoFen-3 mission: a review, *Ieee Geosci. Remote Sens. Mag.*, 13, 79–115, <https://doi.org/10.1109/MGRS.2024.3523924>, 2025.
- Xu, H.: Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, *Int. J. Remote Sens.*, 27, 3025–3033, <https://doi.org/10.1080/01431160600589179>, 2006.
- Zafarmomen, N., Samadi, V., and Borgomeo, E.: Spatiotemporal SWMM-LSTM surrogate modeling for efficient node-level water depth and inflow prediction in urban drainage networks, *Cambridge Prisms: Water*, 4, 4e, 1–35, <https://doi.org/10.1017/wat.2026.10013>, 2026.

- Zhang, J., Zhou, C., Xu, K., and Watanabe, M.: Flood disaster monitoring and evaluation in China, *Global Environmental Change Part B: Environmental Hazards*, 4, 33–43, <https://doi.org/10.3763/ehaz.2002.0404>, 2002.
- 690 Zhang, M., Liu, D., Wang, S., Xiang, H., and Zhang, W.: Multisource remote sensing data-based flood monitoring and crop damage assessment: a case study on the 20 July 2021 extraordinary rainfall event in Henan, China, *Remote Sens.*, 14, 5771, <https://doi.org/10.3390/rs14225771>, 2022.
- Zhao, J., Pelich, R., Hostache, R., Matgen, P., Wagner, W., and Chini, M.: A large-scale 2005–2012 flood map record derived from ENVISAT-ASAR data: United Kingdom as a test case, *Remote Sens. Environ.*, 256, 112338, <https://doi.org/10.1016/j.rse.2021.112338>, 2021.
- 695 Zhao, L., Zhang, Q., Li, Y., Qi, Y., Yuan, X., Liu, J., and Li, H.: China’s Gaofen-3 satellite system and its application and prospect, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, 14, 11019–11028, <https://doi.org/10.1109/JSTARS.2021.3122304>, 2021.