Proceedings of the 4th International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/2026.33085

Flood Information Extraction Based on Sentinel-1 SAR Imagery: A Case Study of Yanshan District, Guilin

Qinru Li[1], Mingyang Song[2*]

[1]Gansu Lintao High School, Dingxi, China
[2]School of Earth System Science, Tianjin University, Tianjin, China
**Corresponding Author. Email: eternal_79838@tju.edu.cn*

Abstract. To solve the problem that optical remote sensing cannot monitor floods properly during cloudy and rainy weather, this study took Sentinel-1 SAR images as the main data source. Using the Google Earth Engine (GEE) cloud platform, this paper combined the Standard Dual-Polarisation Water Index (SDWI) and Random Forest (RF) methods to quickly extract and analyse the flooded area from the severe flood event that hit Yanshan District in Guilin City in June 2024. The results show that both methods performed well in extracting flood information. The SDWI method reached an overall accuracy (OA) of 98.41% and a Kappa coefficient of 0.9590, while the RF method reached an OA of 97.14% and a Kappa coefficient of 0.9190. The flood-inundation area mainly spreads along the banks of major rivers, such as the Lijiang River and the Xiangsi River. When this paper compared the inundation map with population and land-use data, this paper found that the affected zones were mostly cropland and built-up land, which overlapped closely with the most densely populated areas. This led to quite serious impacts. This study confirms that SAR imagery is effective for emergency flood monitoring in karst regions and offers a useful technical reference for fast disaster assessment.

Keywords: microwave remote sensing, flood monitoring, polarised SAR, Dual-polarisation water index, random forest

1. Introduction

Compared with the limitations of human activities and the intensification of global climate change, extreme rainfall events are becoming increasingly frequent. These events have resulted in numerous casualties and enormous economic losses from floods worldwide, posing a serious threat to the stability of society [1-5]. So, compared to dealing with floods, prevention has become particularly urgent, this paper need to conduct research based on actual situations to understand how to avoid difficulties before disasters occur and how to quickly carry out rescue operations during disasters, Accurately tracking areas covered by floods is crucial for making prompt rescue decisions, supporting post disaster recovery, and planning sustainable land use in such situations [4,6].

Its traditional methods, such as checking ground floods or drawing maps by hand, have been punished, which is compared to response time of covered area, They cannot meet immediate monitoring needs of large industries, Remote sensing skills have become main tool for flood

DOI: 10.54254/2755-2721/2026.33085

monitoring and assessment, providing a panoramic view that can quickly gather information and repeatedly observe same area [4,7]. Previously, optical remote sensing imagery was commonly used to identify water bodies. Qiao and his team applied the Normalised Non Water Index (NDWI) to Enhanced Special Topic Imager (ETM+) images and achieved good results in detecting water [8]. Floods often bring prolonged clouds and rain, and optical information is often unusable.

Artificial Aperture Radar (SAR) operates by emitting its own electromagnetic waves. It passes through clouds and even some plants, allowing it to observe floods regardless of the weather. SAR has become an important method for flood monitoring and evaluation research [7]. The Sentinel-1 satellite of the European Space Agency provides free, high-resolution C-band images that cover the entire world and are frequently returned to track changes. The initial stage of SAR flood mapping mainly relied on simple critical values or change records; Chen's team has developed an intelligent method based on the Gamma distribution [9]. In contrast, Yang's team utilises the idea of mirror refraction [3]. These methods are relatively simple and easily susceptible to interference from noise or mixed signals [2]. Recently, machine learning has developed rapidly compared to deep learning, greatly improving way this paper extract material properties in situational descriptions, Sun's team used U-Net and HRNet to segment flood areas in SAR images [6], Acharya's team tested various algorithms and found that Random Forest (RF) had highest accuracy [10], Chen's team achieved high accuracy using Convolutional Neural Networks and DeepLabv3 [11,12].

2. Methods

2.1. Overview of the study area

As shown in Fig. 1, this research uses Yanshan District, Guilin City, Guangxi Zhuang Autonomous Region, China, as the study area. Guilin City is located in the northeastern part of the Guangxi Zhuang Autonomous Region. It is situated in the Lijiang River basin, where karst landforms are well-developed. Guilin has a subtropical monsoon climate. Guilin has abundant rainfall, complex terrain, and a dense network of waterways. Rainstorms and floods occur frequently [13].

Figure 1. Satellite image of Yanshan District

DOI: 10.54254/2755-2721/2026.33085

Yanshan District (Fig. 1) is located in the southern part of Guilin City. It is connected to Lingchuan County in the east, adjacent to Lingui District in the west, borders on Yangshuo County in the south, and shares boundaries with Xiangshan District and Qixing District in the north. Yanshan District is located in low latitudes and receives abundant rainfall, with an annual average of 1894 millimetres. It is a typical karst area. There are the Liangfeng River, the ancient Guiliu Canal and others as the main rivers. The Lijiang River and the Xiangsi River also run through the Yanshan District. The entire region is dotted with numerous stone mountains and hills. There is a plain on the northeast corner of Yanshan Mountain, with an area of stone mountains along the west bank of the Lijiang River on the front side. The highest peak in the Stone Mountains reaches an elevation of 451 meters. There is a vast river-valley basin centred on Zhemu Town, extending southward to Yanshan Mountain and Dabu. Dense karst peak forests surround this basin, while the central area has flat terrain dotted with paddy fields, dry land, hills, rivers, reservoirs, ponds, etc. Most of the terrain is flat, accounting for about 57.6%, and the rest is mostly low-hilly slope [14].

In mid-June 2024, Guilin experienced continuous heavy rainfall, with five major rivers exceeding the warning water stage all along. Two floods exceeded the highest safety water stage around June 19, posing a serious threat to infrastructure and the safety of local people's lives and property [15]. Yanshan District suffered severe damage in this catastrophic flood event. As a result, using the Yanshan District for flood monitoring holds significant practical importance.

2.2. Data source and preprocessing

This study uses multi-source remote sensing data for collaborative analysis. The core data are the European Space Agency's Sentinel-1 SAR imagery (Fig. 2). This paper chose the Ground Range Detection (GRD) product in Interferometric Wide (IW) mode, which includes VV and VH dual-polarisation data. The detailed information can be found in Table 1. Through the GEE, this paper obtained ASCENDING imagery for two periods: May 27, 2024 (before the flood) and June 20, 2024 (after the flood). This paper then performed preprocessing operations, including median composite, radiometric calibration, terrain correction, and speckle filtering (Lee speckle filter).

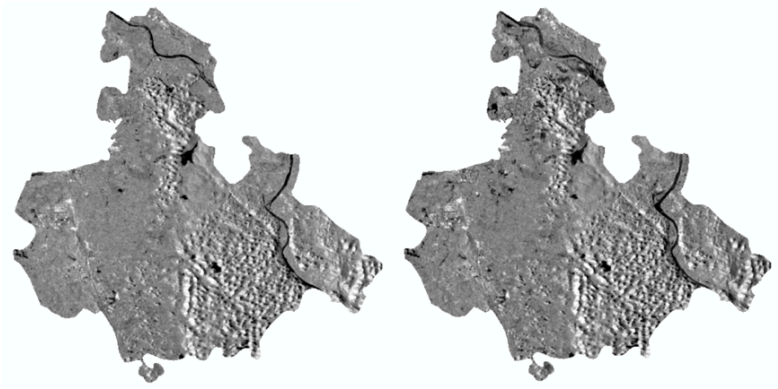

Figure 2. Sentinel-1 SAR images of Yanshan District before (left) and after (right) the flood

Table 1. Sentinel-1 SAR image data

Data Source	Sensor	Acquisition Time	Spatial Resolution	Polarization
Sentinel-1 GRD	C-SAR	2024-05-27 2024-06-20	10m	Dual polarization

DOI: 10.54254/2755-2721/2026.33085

The following is auxiliary data. First, ESA WorldCover 10m v100, which is used to generate and manually supplement sample points automatically. Second, the SRTM (Shuttle Radar Topography Mission) digital elevation model (DEM), which is used for extracting terrain factors such as slope and elevation. Third, WorldPop population data were used to analyse the flood-inundation area and briefly assess the disaster situation. Lastly, the distribution data on land-use types from MAPWORLD are used to confirm land-cover types in flood-affected areas for disaster assessment.

All the above data are set at a spatial resolution of 10m and projected onto the WGS84 geographic coordinate system. Then the image is cropped to the study area boundary for subsequent research.

2.3. Experiment methods

This study used the standard dual-polarisation water index (SDWI) and random forest (RF) methods to extract water.

2.3.1. SDWI method for water extraction

The SDWI method is a threshold-based segmentation approach proposed by Jia et al. for extracting water-body information. Perform band operations on the processed dual-polarised Sentinel-1 data according to equation (1). Pixels with SDWI values greater than 0 are considered water bodies, while those with SDWI values less than 0 are considered non-water bodies. Finally, this paper obtained a binary graph of water body information.

$$SDWI = ln(10 * VV * VH) - 8 \quad (1)$$

According to SAR image characteristics, water bodies generally exhibit specular reflection, resulting in a lower backscattering coefficient. The SDWI method multiplies VV- and VH-polarised images to enhance water features and increase the contrast between water bodies and other land types, thereby extracting water information more accurately.

2.3.2. RF method for water extraction

Random Forest (RF) is an ensemble learning algorithm that uses decision trees as its learning machines, proposed by Breiman in 2001. Its core is to construct a large number of independent decision trees, learn from randomly partitioned training samples, and then integrate the outputs of each tree to obtain the optimal classification results. This study used the random forest algorithm to extract water bodies on the GEE platform. This paper used GEE's built-in RF classifier. The number of decision trees is set to 100. The samples are automatically generated and manually supplemented based on ESA WorldCover. Then they are randomly divided into training and validation sets in a 70%:30 % ratio.

2.3.3. Accuracy evaluation indicators

This study used two indicators, overall accuracy (OA) and the Kappa coefficient, to verify accuracy. The expression is as follows.

DOI: 10.54254/2755-2721/2026.33085

$$OA = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

$$Kappa = \frac{P_0-P_e}{1-P_e} \tag{3}$$

$$P_e = \frac{(FP+TP)*(FN+TP)+(FN+TN)*(FP+TN)}{n^2} \tag{4}$$

In the formula, TP is the number of pixels detected as water bodies that are actually water bodies. TN is the number of pixels detected as non-water bodies and actually non-water bodies. FP is the number of pixels detected as water bodies, but actually non-water bodies. FN is the number of pixels detected as non-water bodies but actually water bodies. P0 is OA, Pe is the expected consistency rate, and n is the total number of pixels.

3. Results

3.1. Accuracy evaluation of water extraction

Using the SDWI and RF methods, the pre-flood water body (Fig. 3) and the post-flood water body range (Fig. 4) in Yanshan District were successfully extracted from Sentinel-1 SAR images. The comparison of accuracy between the two methods is shown in Table 2. It can be seen that the accuracies of the SDWI and RF methods are close, and the results are highly consistent.

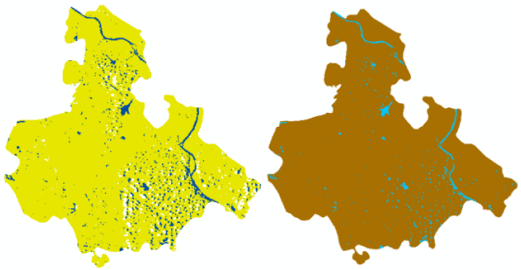

Figure 3. Water extraction results before the flood using the SDWI method (left) and the RF method (right)

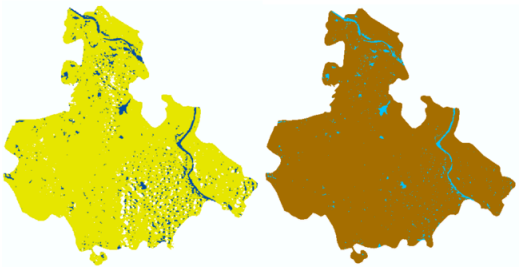

Figure 4. Water extraction results after the flood using the SDWI method (left) and the RF method (right)

DOI: 10.54254/2755-2721/2026.33085

Table 2. Accuracy evaluation results of the SDWI method and the RF method

Accuracy Evaluation Indicators	SDWI	RF
OA(%)	98.41	97.14
Kappa coefficient	0.9590	0.9190

3.2. Flood inundation area extraction

Based on the above results, the inundation range map for this major flood in Yanshan District was obtained using change detection (Fig. 5). The results show that the flood mainly spread along the banks of the Lijiang River, Xiangsi River, and other tributaries. Additionally, the flood also had a large distribution around the water areas near Lijia Village and Qifeng Town.

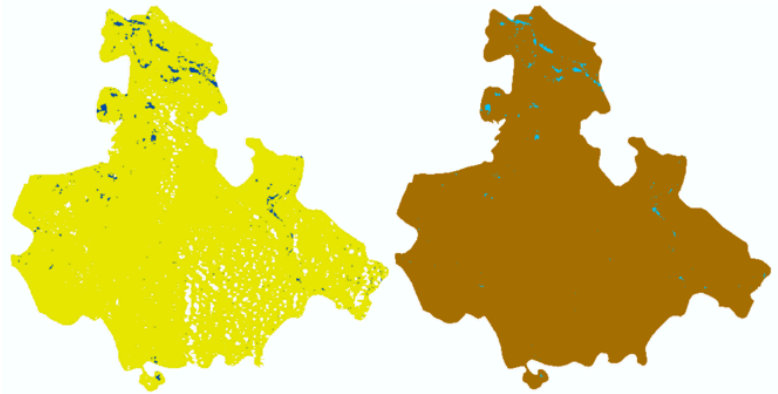

Figure 5. The results of flood inundation area extraction using the SDWI method (left) and the RF method (right)

4. Discussion

Accuracy of two methods is shown in Table 2, overall accuracy (OA) of SDWI method is 98,41%, with a Kappa coefficient of 0,9590, slightly higher than 97,14% and 0,9190 of RF method, This is SDWI multiplied by straight VV and VH polarization bands, highlighting low backscattering characteristics of water, It performs particularly well on flat river valley plains in Yanshan region, where it does not rely heavily on training samples and avoids some common errors, RF method can better handle more complex hilly edges because its numerous decision trees work together to judge scene and filter noise, reducing confusion between shadows and water, water body diagrams in Figures 2 and 3 clearly indicate that these two methods complement each other, SDWI draws clearer boundaries around water bodies, resulting in fewer false alarms generated by RF, These advantages mean that combining these two methods yields more accurate results in undulating and densely mountainous terrain of Yanshan region, This supports conclusion that both SDWI and RF methods efficiently extract flood information with very similar accuracy, indicating that both traditional index based methods and machine learning methods are suitable for this type of work.

Turning to flooded industry in Figure 5, most of flood products originate from Li River, while situation along Si River is different; Compared to lowlands surrounding Qifeng Town, Lijia Village experiences significant flooding, In this case, using DEM for research is more appropriate, DEM research shows that almost all submerged areas are located below an altitude of 50 meters and have a slope of less than 5 °, making them prone to water accumulation, When this paper compare pre

DOI: 10.54254/2755-2721/2026.33085

flood water body in Figure 3 with post flood water body in Figure 4, total water area increases by about 30%, mainly spreading towards valley plain.

Condensed results from these observations are, compared to others, straightforward and consistent in their primary key points: the main submerged areas are located along the Lijiang River, including the banks of the Xiangsi River and around Lijia Village, as well as around Qifeng Town.

By comparing immersion maps in Figure 6, it can be seen that accumulation of land use information and population data shows that about 60% of immersion areas are farmland or built-up areas, These areas completely overlap with most densely populated areas, with WorldPop records showing over 500 people per square kilometer, Thus, number of people affected exceeds 20% of total population of Yanshan District, Not only will it damage farmland and buildings, but floods will also intensify due to difficulty of urban drainage systems, Qifeng Town has a water depth of 1-2 meters and human activities such as surface hardening as shown in initial flood records, situation is closely related to third point in summary: floods mainly cover northern farmland, built-up areas, and some forest patches; Extremely close to population distribution; And it affects many residents.

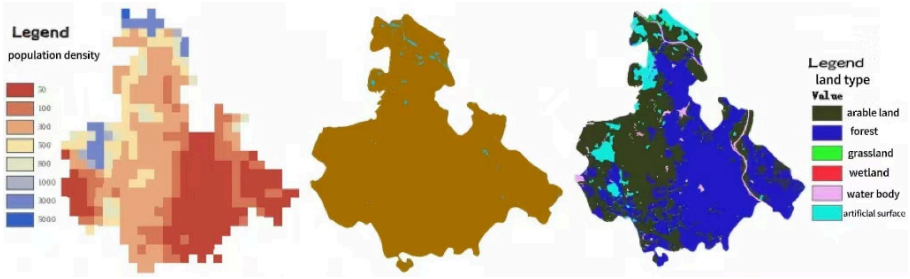

Figure 6. Comparative analysis results of population distribution data (left), flood inundation area (middle), and land use type data (right)

5. Conclusion

This study was based on Sentinel-1 SAR image data. It used SDWI and RF methods to extract water bodies in the Yanshan District of Guilin City. It then further monitored and analysed the flood-inundation area in Yanshan District. For the SDWI method, the OA of the water extraction result is 98.41%, and the Kappa coefficient is 0.9590. For the RF method, the OA of the water extraction result is 97.14%, and the Kappa coefficient is 0.9190. The conclusion drawn from this article is as follows. Firstly, the area of Yanshan District that was submerged by the flood in June 2024 was mainly distributed along the banks of the Lijiang River and Xiangsi River, as well as near Lijia Village and Qifeng Town. Next, both SDWI and RF methods can effectively extract flood water information from Sentinel-1 SAR images. The accuracy of the two methods is close in this case, indicating the feasibility of both traditional index methods and machine learning methods in such applications. Finally, the flooded areas mainly consist of cultivated land, construction land and some forests in the north, which are highly consistent with the population distribution in Yanshan District and affect a large number of people.

This study provides key information support for emergency response and disaster assessment of flood disasters in Yanshan District through remote sensing monitoring. However, the research still has some shortcomings that warrant further exploration and expansion in the following areas. First, collect more remote sensing images and integrate multi-source data, such as optical images, for synergistic analysis to improve the comprehensiveness and reliability of monitoring. Second, adopt deep learning models for comparison and optimisation to improve accuracy and automation further.

DOI: 10.54254/2755-2721/2026.33085

Lastly, combine multiple factors to construct a comprehensive dynamic risk assessment model that incorporates multidimensional indicators to evaluate flood disaster risks with greater granularity.

Authors contribution

All the authors contributed equally, and their names were listed in alphabetical order.

References

[1] Li, X., Lyu, X., & Li, T. (2025). AI and remote sensing in flood control: Applications and future directions. China Flood & Drought Management, 35(9), 1-7, 30.
[2] Gao, H., Chen, B., Sun, H., et al. (2023). Flood detection using SAR satellite imagery: A review of recent advances and future prospects. Journal of Geo-Information Science, 25(10), 1933-1953.
[3] Yang, C., & Chen, D. (1998). Mapping flood inundation extents from spaceborne radar data: A methodological exploration. Journal of Natural Disasters, (3), 46-51.
[4] Xiao, K. (2018). GIS-based flood risk assessment: A case study of Quzhou, Zhejiang Province. Journal of Agricultural Catastrophology, 8(4), 4-7.
[5] Sun, S., Liang, N., Liu, P., et al. (2022). Flood segmentation from Sentinel-1 SAR imagery using deep learning approaches. Modern Surveying and Mapping, 45(6), 12-15.
[6] Zhang, W., Yang, S., Wang, L., et al. (2012). Synthetic aperture radar data for disaster reduction: A review of its potential. Remote Sensing Technology and Application, 27(6), 904-911.
[7] Qiao, C., Luo, J., Sheng, Y., et al. (2012). An adaptive water extraction method from remote sensing image based on NDWI. Journal of the Indian Society of Remote Sensing, 40(3), 421-433.
[8] Chen, L., Liu, Z., & Zhang, H. (2014). Water detection in SAR imagery using scattering characteristics of water bodies. Remote Sensing Technology and Application, 29(6), 963-969.
[9] Acharya, T. D., Subedi, A., & Lee, D. H. (2019). Evaluation of machine learning algorithms for surface water extraction in a Landsat 8 scene of Nepal. Sensors, 19(12).
[10] Guilin Municipal People's Government Office. (2026, January 22). Overview of Guilin. Guilin Municipal People's Government. https: //www.guilin.gov.cn/glyx/glgk/
[11] Peng, G., & Tang, T. (2024). Early preparation, rapid response, and strong rescue: Safeguarding lives and property during the June 19 extreme flood in Guilin, Guangxi. Disaster Reduction in China, (19), 22-24.
[12] Jia, S., Xue, D., Li, C., et al. (2019). Extracting water bodies from Sentinel-1 data: A methodological study. Yangtze River, 50(2), 213-217.
[13] He, R. (2025). Flood extent extraction and analysis in the Liaohe Plain in 2022 using machine learning. Geospatial Information, 23(11), 70-73.
[14] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.
[15] Wu, C., Lai, G., Huang, H., et al. (2026). Remote sensing inversion of local topography around Poyang Lake using a random forest hybrid model. Journal of Lake Sciences, 1-15.