STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, July 2026, pp 619–640.

Hybrid SAR-Optical Remote Sensing for Flood Inundation Mapping: Feature Contribution Analysis in a Wetland Environment

Desy Ika Puspitasari[1,2,*], Edi Noersasongko[1], Purwanto [1], Moch. Arief Soeleman[1], Catur Supriyanto[1]

[1] *Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, Indonesia*
[2] *Faculty of Information Technology, Universitas Islam Kalimantan, Muhammad Arsyad Al Banjari, Banjarmasin, Indonesia*

Abstract Flooding is a recurrent problem in Banjar Regency due to its low-lying wetland topography, high rainfall, and river sedimentation. Nonetheless, precise flood mapping continues to be a challenge due to the interference of cloud cover on optical imaging and spectral ambiguity in diverse wetland ecosystems. To address these limitations, this study proposes a hybrid SAR–optical method that combines Sentinel-1 backscatter with Sentinel-2 water indices (FWEI and AWEI) using pixel-level feature stacking and feature contribution analysis. Four machine learning classifiers—Random Forest, Logistic Regression, Support Vector Machine (SVM), and XGBoost—were evaluated using a spatially independent validation scheme to ensure robust generalization. The results indicate that AWEI surpasses individual features in water discrimination, whereas the combination of SAR and AWEI consistently enhances classification stability and spatial coherence. Among the assessed models, SVM and RF exhibit similar performance, with SVM attaining the optimal classification balance (OA = 0.98, Kappa = 0.97, F1 = 0.98) for water detection, while the difference lacks statistical significance. Flood inundation maps were derived through change detection between independently classified pre- and post-flood water maps. The results indicate that multi-sensor integration enhances the delineation of flood-affected areas, particularly in complex wetland environments. However, the study also highlights that the reliability of flood inundation mapping is inherently dependent on the accuracy of pre- and post-flood water classification as well as the temporal consistency of the input imagery. Furthermore, feature contribution analysis combined with spatial validation reveals that AWEI provided the strongest standalone classification performance, while SAR contributed complementary structural information that improved classification robustness. These findings demonstrate that spatial validation effectively mitigates autocorrelation bias, leading to more reliable and generalizable water classification performance for flood inundation analysis.

Keywords Hybrid SAR–optical, Feature contribution, Spectral indices, Spatial validation, Wetland environment, Remote Sensing, Machine Learning

DOI: 10.19139/soic-2310-5070-4067

1. Introduction

Floods are among the most destructive hydrometeorological disasters, causing substantial damage to infrastructure, land-use systems, and human livelihoods [1]. Accurate flood monitoring and reliable flood inundation mapping are therefore crucial for offering emergency response and post-disaster recovery planning, agricultural water management, crop damage assessment, and effective drainage and mitigation strategies [2], [3]. Banjar Regency experiences a tropical monsoon climate, with a pronounced rainy season from October to May, peaking between December and March. These conditions frequently trigger seasonal flooding in low-lying areas. In early 2023, floods affected 99 villages, damaging approximately 17,000 dwellings and impacting nearly 66,000 residents [4, 5, 6], highlighting the critical necessity for precise and prompt flood inundation mapping. The region's

---

[*] Correspondence to: Desy Ika Puspitasari (Email: p41202300071@mhs.dinus.ac.id). Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, Central Java Province, Indonesia (50131).

ISSN 2310-5070 (online) ISSN 2311-004X (print)

vulnerability is intensified by its low-lying wetland topography, the presence of the Martapura River and its tributaries, sedimentation from upstream activities, and land-use changes that increase surface runoff [7, 8].

Recent literature indicates that optical remote sensing, especially Sentinel-2, has predominated in flood detection owing to its extensive spectral information [9, 10]. Numerous spectral water indices have been formulated to improve the delineation of water bodies, including the Normalized Difference Water Index (NDWI), Modified Normalized Difference Water Index (MNDWI), Automated Water Extraction Index (AWEI), and Flood/Water Extraction Index (FWEI). The development of more robust indices, such as FWEI and AWEI, offers a potential solution to suppress noise in heterogeneous areas [11]. However, automated thresholding methods like the Otsu method still rely on imagery with high pixel variability [12].

Complementarily, Sentinel-1 Synthetic Aperture Radar (SAR) offers essential cloud-penetrating capability during the rainy season [13]. However, similar backscatter from still water surfaces and other smooth wetland surfaces often hampers single-SAR approaches [14]. The closest Sentinel-1 and Sentinel-2 acquisitions to the documented flood peak date were selected for pre- and post-flood analysis to document the early 2023 flood. However, optical data during the rainy season were largely affected by cloud cover, limiting the consistency of Sentinel-2 surface observations in several areas. Therefore, Sentinel-1 SAR data played an important complementary role due to its cloud-penetrating capability, particularly for post-flood surface characterization under persistent cloudy conditions.

Despite the growing adoption of multi-sensor SAR-optical integration, two key research gaps remain underexplored. First, most data fusion studies tend to use direct fusion techniques without systematically conducting feature contribution analysis to quantify the relative contributions of each spectral parameter and backscatter [15]. Second, there is a problem of ”over-optimism” in assessing model accuracy due to the use of randomly selected validation data, which often ignores spatial autocorrelation and leads to failure to generalize the model to new areas [16]. Without spatially independent validation, the reliability of models in real-world conditions is uncertain due to spatial autocorrelation, which may lead to overly optimistic performance evaluations [17].

This paper proposes a hybrid machine learning approach that combines Sentinel-1 backscatter data with Sentinel-2 water indices (FWEI and AWEI) using a pixel-level feature stacking methodology to tackle these problems. Four machine learning techniques were chosen: Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), and XGBoost, to evaluate water classification performance for flood inundation analysis in wetland environments. These algorithms can manage high-dimensional data and elucidate intricate nonlinear correlations resulting from the amalgamation of SAR backscatter properties and optical spectral indices. Random Forest and XGBoost were chosen for their resilience to data noise and capacity to manage complex characteristics [18, 19, 20]. Support Vector Machines (SVM) are proficient in class separation inside high-dimensional feature spaces, even when training datasets are limited [21]. At the same time, LR serves as a baseline model to deliver effective probabilistic interpretation in the binary classification task, distinguishing water and non-water classes [22].

This study's main contribution is the application of feature contribution analysis to investigate the individual and synergistic impacts of SAR and optical sensors on classification performance. Additionally, we utilize a spatially independent validation approach to reduce autocorrelation bias, offering a more dependable evaluation of the model's generalizability than traditional thresholding techniques. The primary aims of this study are (1) to create a hybrid SAR-optical approach utilizing pixel-wise feature stacking; (2) to assess the efficacy of different machine learning algorithms in classifying water and non-water for flood inundation analysis in wetland ecosystems; and (3) to analyze feature contributions and conduct spatial validation to evaluate the robustness and generalization performance of the proposed model. This study enhances the flood mapping literature by offering a statistically proven and technically generalizable methodology applicable to regions with analogous geographic characteristics.

2. Materials and Methods

2.1. Study Area

The research region is in the province of South Kalimantan, Indonesia, specifically in Banjar Regency (Fig. 1). Banjar Regency has 20 sub-districts and covers about 4,668 km². It is located between 2°49’55”-3°43’38” S and 114°30’20”-115°33’37” E. The area has many low-lying wetlands and freshwater marshes that are quite susceptible to changes in water levels. It has a tropical rainforest climate, which means it rains a lot every year. The monsoon season, which lasts from November to March, is when flooding happens the most.

Distinguishing between permanent water bodies, saturated soils, and ephemeral flood inundation using remote sensing data is difficult due to the flat terrain and dense vegetation. The Barito, Martapura, and Riam Kanan rivers are significant contributors to the hydrological system. They induce seasonal inundation during periods of intense rainfall. The natural characteristics of the locality make it an excellent site for evaluating methodologies that employ multiple sensors to detect floods, including water-level sensors, rainfall gauges, and remote sensing, for monitoring and forecasting flood events.

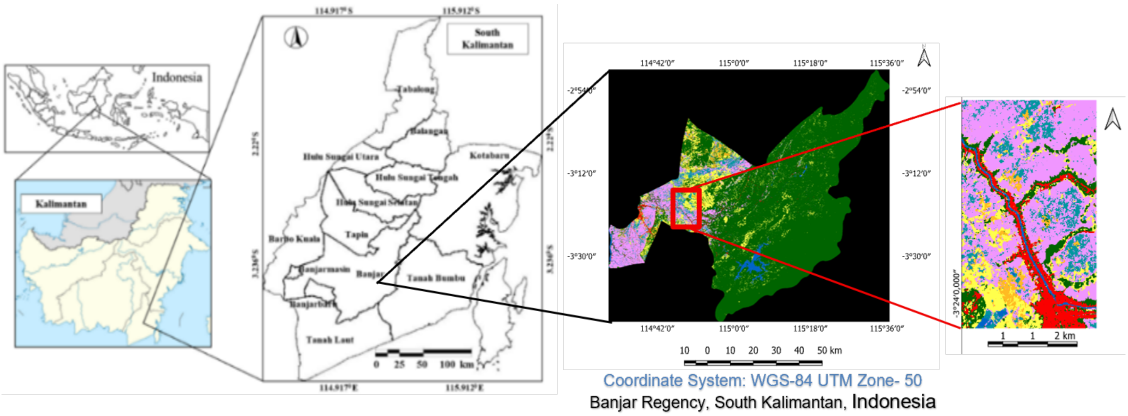


Figure 1. Area of Interest, Banjar Regency, South Kalimantan, Indonesia

2.2. Proposed Method

The proposed hybrid SAR–optical flood inundation mapping is structured as a modular, multi-stage processing pipeline, as illustrated in Figure 2. This method integrates heterogeneous remote sensing data sources, including Sentinel-1 SAR backscatter (VV and VH polarizations) and Sentinel-2 multispectral imagery, to exploit complementary radar and optical information. Spectral water indices, namely the Floating Water Extraction Index (FWEI) and the Automated Water Extraction Index (AWEI), are derived from Sentinel-2 data, while ESA World Cover data are used to support reference-guided sample preparation and post-processing.

In the preprocessing stage, Sentinel-1 SAR data undergo radiometric calibration, speckle filtering, terrain correction, and spatial resampling to ensure radiometric stability and geometric consistency. A Refined Lee speckle filter with a 3×3 kernel size was applied to Sentinel-1 SAR data to reduce speckle noise while preserving spatial details. Sentinel-2 imagery is subjected to atmospheric correction, cloud masking, and spatial resampling to achieve co-registration and spatial alignment with the SAR data.

To maintain spatial consistency, Sentinel-2 imagery was resampled to 10 m by bilinear interpolation, while Sentinel-1 SAR data were resampled using the nearest neighbor method to preserve backscatter characteristics. This harmonization step enables pixel-level feature integration across sensors. During feature extraction, radar backscatter features (VV and VH) are combined with optical water indices (FWEI and AWEI) to construct a unified feature space. A feature-level contribution technique is employed to assess the individual and collective

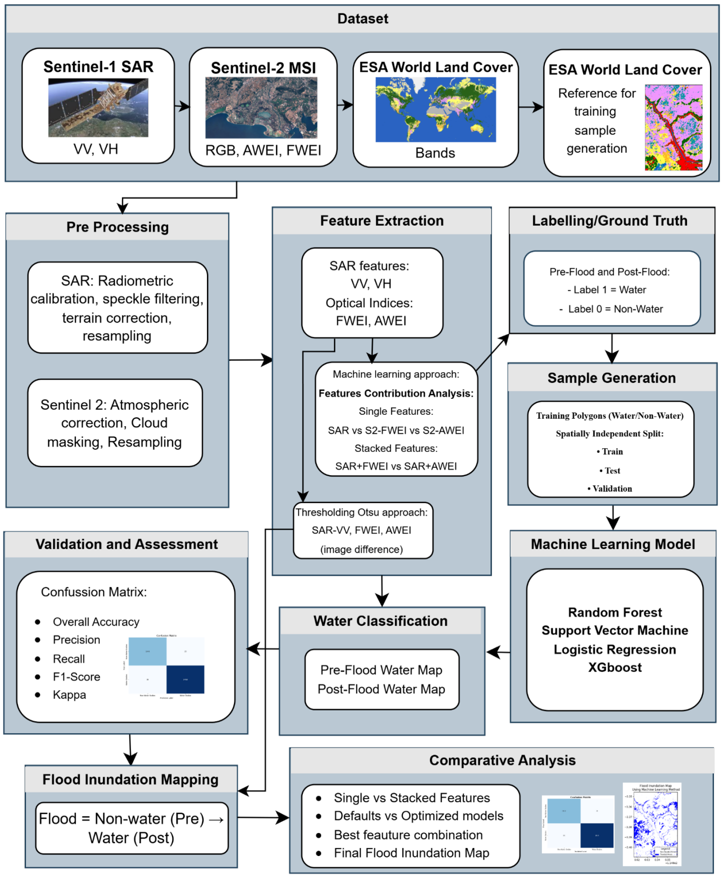


Figure 2. Workflow of the Hybrid SAR–Optical for Flood Inundation Mapping

contributions of each feature set. Accordingly, experiments are conducted using single-source inputs (SAR-only, FWEI-only, and AWEI-only) as well as multi-sensor combinations (SAR + FWEI and SAR + AWEI).

Supervised classification is executed utilizing several machine learning techniques, such as Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), and Extreme Gradient Boosting (XGBoost). To mitigate spatial autocorrelation bias and avoid overly optimistic accuracy estimates, training samples are manually digitized from pre- and post-flood imagery and partitioned into spatially independent training, validation, and testing subsets.

Simultaneously, a conventional threshold-based baseline approach was implemented using Otsu's algorithm applied to SAR backscatter and optical index changes derived from pre- and post-flood image differencing. The Otsu method was included as a reference baseline to provide comparative insights into traditional unsupervised flood detection approaches. Within the supervised machine learning method, binary water/non-water classification maps are generated before flood delineation, whereas the Otsu-based method directly identifies inundated areas from image change information without explicit intermediate water classification.

In the machine learning method, water bodies were first classified independently for pre- and post-flood conditions. Flood inundation was subsequently derived through temporal differencing, where newly flooded pixels were identified as transitions from non-water (pre-flood) to water (post-flood). Model performance was evaluated using confusion matrix-derived metrics, including overall accuracy, precision, recall, F1-score, and Cohen's kappa coefficient.

2.3. Dataset

Sentinel-1 imagery is obtained from two polar-orbiting satellites, S1A and S1B, which are outfitted with C-band Synthetic Aperture Radar (SAR) sensors operate in Interferometric Wide Swath (IW) mode [14]. These satellites provide dual-polarization backscatter data, including vertical–vertical (VV) and vertical–horizontal (VH), with a spatial resolution of approximately 10 m and a combined revisit cycle of six days. SAR is capable of capturing backscatter variations associated with surface water and inundation, even under vegetation cover [23]. The Copernicus Sentinel-2 mission (S2A/S2B) delivers multispectral imagery including 13 spectral bands, with spatial resolutions varying from 10 m to 60 m [10]. To ensure spatial consistency and enable pixel-level feature stacking, all datasets were resampled and harmonized to a common spatial resolution of 10 m before feature extraction and model training. The acquisition windows for each sensor are summarized in Table 1.

Table 1. The Acquisition Time of Satellite Imagery

Flood Period	Acquisition Period	Data Source
Pre-Flood	01/06/22 – 31/10/22	https://developers.google.com/earth-engine/datasets
Post-Flood	01/11/22 – 31/05/23	https://developers.google.com/earth-engine/datasets

2.4. Pre-Processing

Sentinel-1 imagery covering the area of interest (AOI) was filtered according to the designated pre-flood (June–October 2022) and post-flood (November 2022–May 2023) acquisition periods and separated into ascending and descending orbits. Backscatter values below -30 dB were masked to reduce edge noise effects. Subsequently, VV backscatter composites were averaged across orbits and clipped to the AOI. The resulting mosaics were visualized within a range of -25 to 5 dB to improve interpretability before feature extraction. Sentinel-2 Level-2A surface reflectance imagery pre-processed using Sen2Cor was acquired through Google Earth Engine (GEE). Scenes within each acquisition period were filtered using a cloud-cover threshold of less than 60%, while cloud and cirrus contamination were removed using the QA60 mask. Reflectance values were converted to Bottom-of-Atmosphere (BOA) reflectance by dividing the digital number values by 10,000. Median composites representing pre-flood and post-flood conditions were subsequently generated and clipped to the AOI. Unlike SAR backscatter data, Sentinel-2 optical imagery was visualized using reflectance values scaled between 0 and 0.3 to enhance the representation of land and water surface characteristics before feature extraction.

2.5. Sampling

Throughout the study region, 50 polygons representing water (class 1) and non-water (class 0) classes were digitized and spatially distributed across heterogeneous wetland conditions. For the pre-flood dataset, the samples consisted of 25 water polygons and 25 non-water polygons, while the post-flood dataset contained 26 water polygons and 24 non-water polygons. These polygons were partitioned into 30 training, 10 validation, and 10 testing polygons using a spatially independent split. The subsets were geographically separated to minimize spatial autocorrelation and reduce optimistic bias in performance estimation. The same polygon partitions were consistently applied across all feature configurations, including SAR-only, optical index only (FWEI, AWEI), and stacked SAR-optical experiments. The labelled polygons were implemented on both pre-flood and post-flood raster datasets to train and evaluate the water/non-water classification models. This process generated independent pre-flood and post-flood water classification maps, which were subsequently used to derive flood inundation extent through pre-post classification differencing. The extracted pixel samples showed relatively balanced class distributions across all dataset partitions. The training set consisted of 3,498 non-water pixels and 4,426 water pixels, while the validation set contained 1,757 non-water and 1,654 water pixels. The testing set included 5,476 non-water and 4,188 water pixels.

2.6. Evaluation

The evaluation of model performance was conducted using spatially independent validation and test datasets obtained from the established polygon partitions. The validation subset facilitated model selection, whereas the independent test subset was designated for final performance assessment to ensure an unbiased evaluation. For the final flood inundation evaluation, an additional set of 200 spatially independent validation points was manually interpreted from pre- and post-flood imagery. The validation samples were not used during model training and were collected independently for the final flood inundation assessment.

Performance evaluation was assessed using metrics derived from the confusion matrix, including overall accuracy (OA), precision, recall, F1 score, and Cohen's Kappa coefficient. The F1-score provides a balanced measure of precision and recall, while the Kappa coefficient assesses categorization concordance beyond chance. These measures jointly offer a thorough evaluation of the model's capacity to differentiate between water and non-water pixels across heterogeneous wetland environments.

2.7. Thresholding Otsu

Otsu's method [24] automatically determines an optimal threshold by maximizing the between-class variance while minimizing the within-class variance of a grayscale histogram. The method is particularly suitable for bimodal histograms, such as those representing flooded and non-flooded pixel distributions [25]. In this study, Otsu thresholding was applied to SAR backscatter and optical index change information derived from pre- and post-flood imagery to generate baseline flood inundation maps. The resulting binary maps classify flooded areas as "1" and non-flooded areas as "0", which were subsequently used as baseline results for comparison with the proposed machine learning-based method. Mathematically, the optimal threshold T is selected by maximizing the between-class variance:

$$\sigma_b^2(T) = P_{WB}(T)\, P_{NWB}(T)\, [\mu_{WB}(T) - \mu_{NWB}(T)]^2 \tag{1}$$

where P_{WB} and P_{NWB} denote the class probabilities, while $\mu_{WB}(T)$ and $\mu_{NWB}(T)$ represent the corresponding class means. The optimal threshold T is obtained by maximizing $\sigma_b^2(T)$.

2.8. Water Extraction Indices

The Automatic Water Extraction Index (AWEI) is designed to improve the discrimination between water and non-water surfaces across diverse environmental conditions. It employs multiple spectral bands, including the green, near-infrared (NIR), shortwave infrared 1 (SWIR1), and shortwave infrared 2 (SWIR2) bands, to stabilize the threshold value used for separating water and non-water pixels. Pixels with index values below 0 are classified as non-water, whereas values above 0 indicate water.

Feyisa et al. (2014) reported that AWEI$_{nsh}$ achieved an overall accuracy of 98% in a study conducted in Denmark using shadow-free imagery, with both user's and producer's accuracies reaching approximately 92%. Previous studies have also demonstrated that AWEI performs reliably for surface water monitoring in Google Earth Engine environments [26]. In addition, AWEI$_{nsh}$ has shown effectiveness in detecting rivers and small water bodies [27], making it particularly suitable for wetland-dominated regions such as the study area. Feyisa et al. (2014) further stated that AWEI$_{nsh}$ [Eq. (1)] is especially effective in areas characterized by dark surfaces. Therefore, this study employed both AWEI$_{nsh}$ and FWEI [11] for water body mapping evaluation:

$$\text{AWEI}_{nsh} = 4(B3 - B11) - (0.25 \times B8 + 2.75 \times B12) \tag{2}$$

$$\text{FWEI} = \frac{M - B8}{M + B8} \tag{3}$$

where

$$M = \frac{B2 + B3 + B4}{3}$$

Equations (2)–(3) are used to calculate water indices based on Sentinel-2 spectral bands.

- B3 represents the green band (0.56 μm), which enhances water reflectance and improves the discrimination of water bodies from surrounding land surfaces.
- B8 denotes the Near Infrared (NIR) band (0.84 μm), which is strongly absorbed by water and is therefore useful for separating water from vegetation.
- B11 represents the Short-Wave Infrared 1 (SWIR1) band (1.61 μm), which is sensitive to moisture content and helps suppress non-water features such as soil and built-up areas.
- B12 denotes the Short-Wave Infrared 2 (SWIR2) band (2.19 μm), which further enhances the separation between water and non-water surfaces, particularly in complex wetland environments.

These spectral characteristics enable effective discrimination between water and non-water surfaces under varying environmental conditions.

2.9. Evaluation Metrics

Accuracy assessment in binary classification tasks, such as water body detection, is fundamentally based on confusion matrix analysis, which summarizes model prediction results by comparing them with reference data (ground truth). The evaluation metrics employed in this study include overall accuracy (OA), recall, precision, F1-score, and Cohen's Kappa coefficient. These metrics collectively provide complementary perspectives for evaluating classification performance. The mathematical formulations of the evaluation metrics are presented in Eqs. (4)–(10):

$$OA = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{7}$$

$$\rho_o = \frac{TP + TN}{N} \tag{8}$$

$$\rho_e = \frac{(TP + FP)(TP + FN) + (FN + TN)(FP + TN)}{N^2} \tag{9}$$

$$Kappa = \frac{\rho_o - \rho_e}{1 - \rho_e} \tag{10}$$

where

$$N = TP + FP + TN + FN$$

True Positives (TP) denote water pixels correctly identified as water, while True Negatives (TN) represent non-water pixels correctly identified as non-water. False Positives (FP) refer to non-water pixels incorrectly classified as water, whereas False Negatives (FN) indicate water pixels misclassified as non-water. Overall Accuracy (OA) measures the proportion of correctly classified pixels relative to the total number of pixels. Recall (sensitivity) evaluates the ability of the model to correctly identify actual water pixels, whereas precision indicates the proportion of predicted water pixels that are truly water. The F1-score represents the harmonic mean of precision and recall, providing a balanced evaluation under imbalanced class distributions. Meanwhile, the Kappa coefficient (κ) measures the agreement between predicted classifications and reference data while accounting for agreement occurring by chance.

3. Results and Discussion

3.1. Visual Comparison of Pre-Flood and Post-Flood

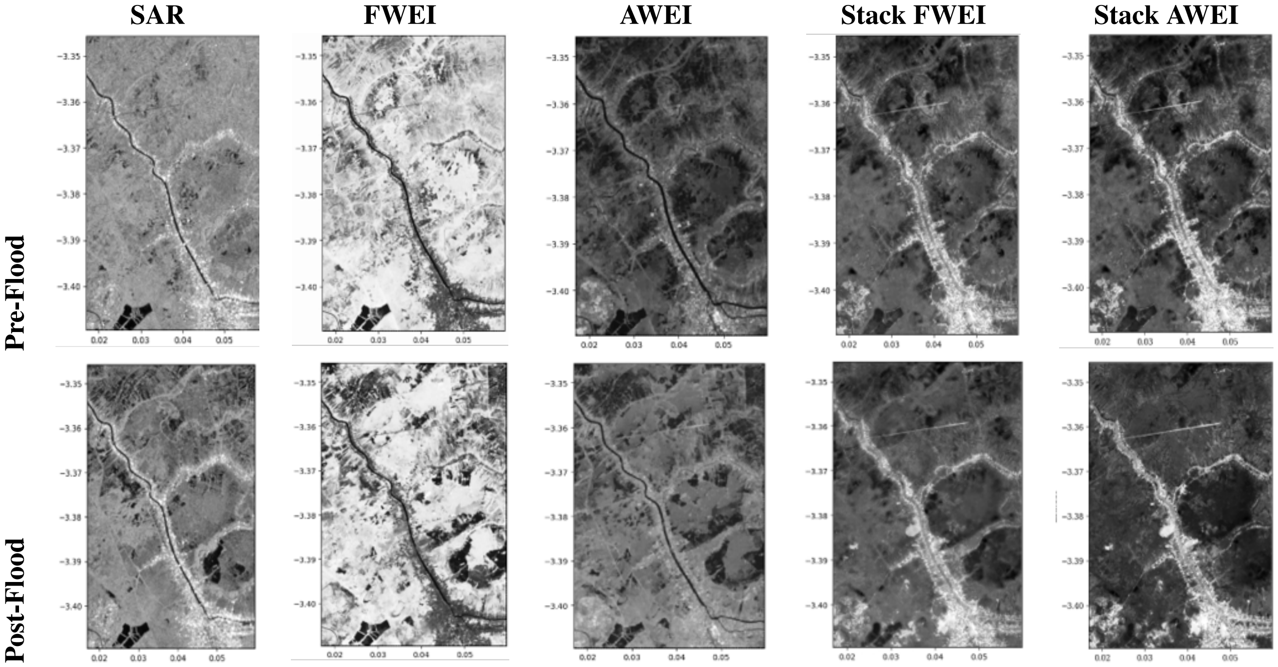
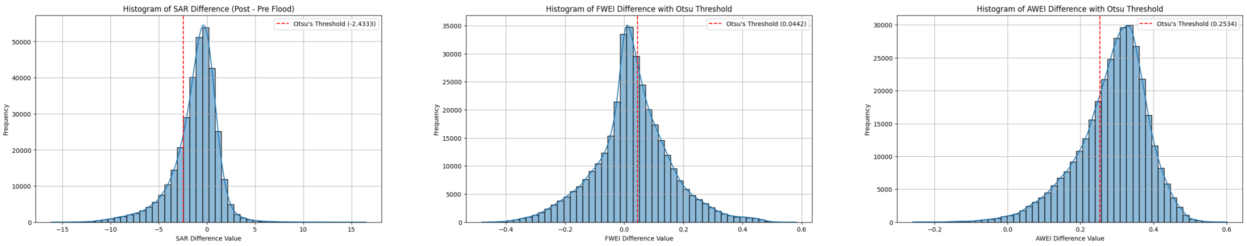


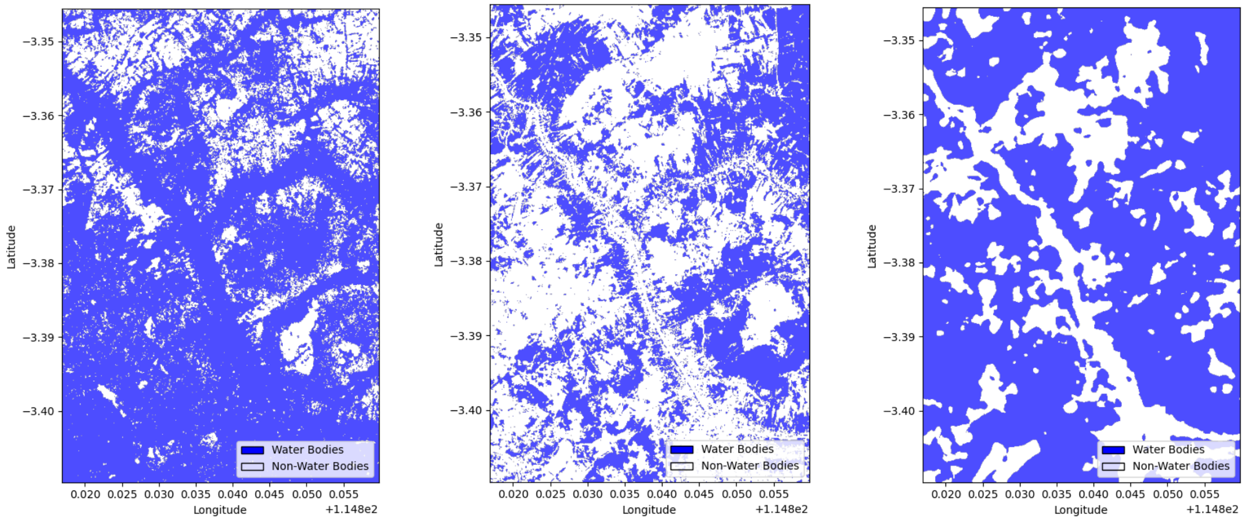
Figure 3. Visual Comparison of Pre-Flood and Post-Flood Imagery Derived from SAR and Optical Water Indices

Figure 3 presents a visual comparison between pre-flood and post-flood conditions using SAR backscatter imagery, optical water indices (FWEI and AWEI), and integrated SAR–optical features. In the SAR imagery, river channels and permanent water bodies appear darker than surrounding land surfaces due to the specular reflection effect of smooth water surfaces, which reduces radar backscatter intensity. The post-flood observations reveal a noticeable expansion of dark regions along river corridors and adjacent floodplains, indicating the presence of newly inundated areas. The study area is dominated by lowland wetlands and freshwater marshes that experience substantial seasonal fluctuations in water levels. During flood events, increased rainfall and river discharge frequently inundate surrounding floodplain regions, resulting in temporary expansion of surface water. In SAR imagery, inundated surfaces generally exhibit lower backscatter values because smooth water surfaces reflect radar signals away from the sensor. However, wetland environments characterized by saturated soils and dense vegetation may also produce low backscatter responses, creating ambiguity between flooded and non-flooded areas when relying solely on SAR data. Optical water indices, including FWEI and AWEI, improve the discrimination between water and surrounding land surfaces by emphasizing spectral absorption characteristics associated with water bodies. These indices enable clearer identification of inundated regions in post-flood imagery. Furthermore, the integration of SAR backscatter and optical water indices (SAR–FWEI and SAR–AWEI) enhances flood delineation performance by combining complementary radar and spectral information. The integrated features reveal more distinct inundation patterns along riverbanks and floodplain areas compared to single-source observations. These observations demonstrate the potential of combining SAR backscatter and optical water indices for flood-related water detection. Nevertheless, reliable flood delineation still requires an appropriate classification or thresholding approach to separate water and non-water surfaces from the extracted features. Therefore, the following section evaluates the effectiveness of the Otsu thresholding method for preliminary flood inundation mapping.

3.2. *Water Body Extraction using Otsu Threshold*



(a) SAR Otsu Histogram (b) FWEI Otsu Histogram (c) AWEI Otsu Histogram



(d) SAR Water Mask (e) FWEI Water Mask (f) AWEI Water Mask

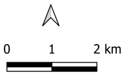


Figure 4. Water Body Extraction using Otsu Thresholding

Figure 4 presents the Otsu thresholding results for SAR backscatter and optical water indices (FWEI and AWEI). The first row shows histogram-based threshold selection, while the second row presents the corresponding binary water masks. Pixels with values above the threshold were classified as water bodies, whereas pixels below the threshold were categorized as non-water bodies. The SAR-based extraction, using a threshold value of -2.4333, produced a dense and spatially fragmented distribution of water pixels across the study area. This pattern indicates that SAR backscatter is highly sensitive to variations in surface roughness and moisture conditions. Consequently, the detected water class includes not only open water surfaces but also saturated soils and vegetated wetland areas.

In contrast, the FWEI-based extraction (threshold = 0.0442) generated a more structured representation of water bodies due to its capability to enhance water-related spectral responses while suppressing background noise. Meanwhile, the AWEI-based extraction (threshold = 0.2534) produced more extensive water patches, suggesting stronger suppression of non-water signals such as vegetation and built-up surfaces, although with a tendency to overestimate flood extent. These findings demonstrate that SAR backscatter, FWEI, and AWEI respond differently to water-related surface characteristics and environmental conditions. Therefore, integrating SAR backscatter information with optical water indices within a machine learning method is expected to improve flood detection reliability and reduce classification ambiguity.

3.3. *Spatial Generalization Evaluation*

To evaluate the spatial generalization capability of the proposed method, the training, validation, and testing samples were spatially distributed across geographically diverse regions within the study area, as illustrated in Figure 5. To ensure strict separation between training and evaluation data, sampling polygons representing water and non-water classes were independently delineated without spatial overlap. This strategy enabled the model to be evaluated on previously unseen regions, thereby providing a more realistic assessment of spatial generalization performance.

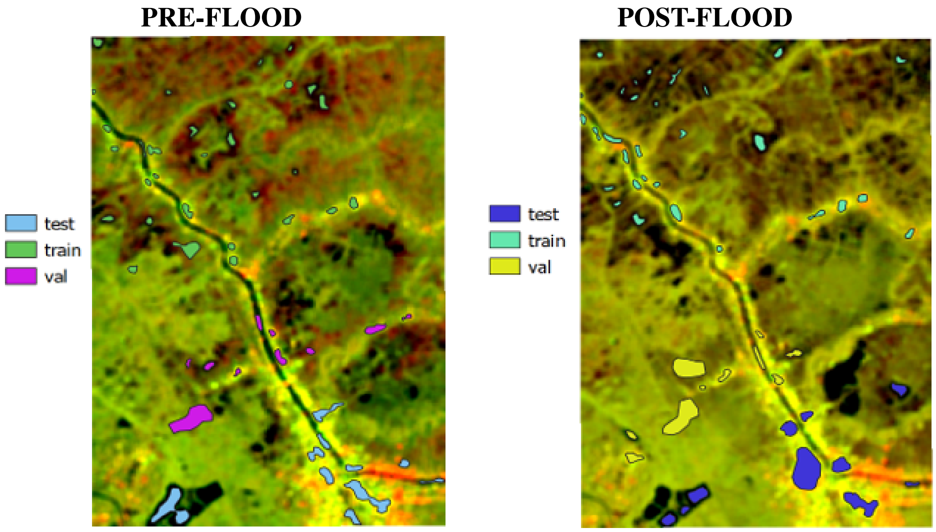


Figure 5. Spatial Distribution of Training, Validation, and Testing Samples in Pre-Flood and Post-Flood Conditions

The validation samples were distributed across multiple landscape characteristics, including river channels, floodplains, wetlands, and surrounding land-cover types under both pre-flood and post-flood conditions. These heterogeneous environmental conditions substantially influence surface reflectance and backscatter characteristics. By spatially separating the evaluation regions from the training regions, the proposed method minimizes spatial autocorrelation bias and provides a more robust assessment of model performance under complex wetland environments.

3.4. *Hyperparameter Tuning Results*

To optimize classification performance, hyperparameter tuning was conducted for each machine learning algorithm using predefined search spaces. The optimal hyperparameter combinations were selected based on validation performance to improve model robustness and generalization capability under heterogeneous wetland conditions. The tuning process aimed to identify parameter settings that balance classification accuracy and model stability while reducing the risk of overfitting. Table 2 summarizes the hyperparameter search space and the best-performing configuration obtained for each machine learning model.

Table 2 presents the hyperparameter search space and the optimal configuration obtained for each machine learning model. The search space defines the range of candidate parameter values explored during the tuning process, whereas the optimal hyperparameters represent the configuration that achieved the best validation performance. For the XGBoost model, the tuning process explored variations in the number of estimators, maximum tree depth, and learning rate. The optimal configuration selected a low learning rate (0.01), shallow tree depth (3), and 100 estimators, indicating a conservative learning strategy that helps reduce overfitting while maintaining stable model convergence.

For the Random Forest model, the number of trees, maximum depth, and minimum samples per leaf were optimized. The selected configuration (max depth $= 10$, min samples leaf $= 4$, and n estimators $= 100$)

Table 2. Hyperparameter Configuration

Model	**Search Space**	**Best Hyperparameters**
XGBoost	{’n_estimators’: [100, 200, 300], ’max_depth’: [3, 5, 7], ’learning_rate’: [0.01, 0.1, 0.2]}	{’learning_rate’: 0.01, ’max_depth’: 3, ’n_estimators’: 100}
Random Forest	{’n_estimators’: [100, 200, 300], ’max_depth’: [10, 20, 30], ’min_samples_leaf’: [1, 2, 4]}	{’max_depth’: 10, ’min_samples_leaf’: 4, ’n_estimators’: 100}
SVM	{’C’: [0.1, 1, 10], ’kernel’: [’linear’, ’rbf’], ’gamma’: [’scale’, ’auto’]}	{’C’: 1, ’gamma’: ’auto’, ’kernel’: ’rbf’}
Logistic Regression	{’C’: [0.1, 1, 10], ’penalty’: [’l1’, ’l2’]}	{’C’: 0.1, ’penalty’: ’l1’}

suggests that limiting tree complexity and increasing the minimum number of samples per leaf improved model generalization by reducing variance and preventing overly complex trees.

For the SVM model, the regularization parameter (C), kernel type, and gamma value were optimized. The best-performing configuration ($C = 1$, kernel = RBF, gamma = auto) indicates that a nonlinear decision boundary was more suitable for capturing complex relationships between SAR backscatter and optical spectral features.

Logistic Regression was tuned using different regularization strengths and penalty types. The optimal configuration ($C = 0.1$, penalty = L1) indicates that stronger regularization improved model interpretability and reduced overfitting by encouraging sparse model coefficients.

Overall, the hyperparameter tuning results demonstrate that appropriate parameter selection plays an important role in improving classification robustness and generalization performance across heterogeneous wetland environments.

All machine learning experiments were implemented in Python 3.12.13 using NumPy 2.0.2, Pandas 2.2.2, Scikit-learn 1.6.1, Rasterio 1.5.0, GeoPandas 1.1.3, XGBoost 3.2.0, and Statsmodels 0.14.6. The implementation scripts and spatial sampling polygons are publicly available at: `https://github.com/desyika-source/SAR-OpticalFloodMapping`.

3.5. Accuracy Comparison of Machine Learning Models

Tables 3–5 summarize the classification performance of the evaluated machine learning models using different feature combinations, including SAR backscatter, optical water indices (S2-FWEI and S2-AWEI), and integrated SAR–optical features. Model performance was assessed using overall accuracy (OA), Cohen’s Kappa coefficient, F1-score, precision, and recall.

Table 3. Accuracy Assessment Results for Water Classification

Model	**RF**			**SVM**			**LR**			**XGB**		
	OA	**Kappa**	**F1**	**OA**	**Kappa**	**F1**	**OA**	**Kappa**	**F1**	**OA**	**Kappa**	**F1**
SAR	0.96	0.91	0.95	0.84	0.60	0.74	0.97	0.93	0.97	0.95	0.89	0.94
S2-FWEI	0.74	0.48	0.74	0.81	0.63	0.80	0.81	0.62	0.80	0.81	0.62	0.81
S2-AWEI	**0.97**	**0.93**	**0.97**	**0.98**	0.96	**0.98**	**0.98**	**0.95**	**0.98**	**0.98**	**0.95**	**0.98**
SAR+FWEI	0.92	0.85	0.92	0.93	0.86	0.92	0.92	0.83	0.87	0.96	0.86	0.93
SAR+AWEI	0.96	0.90	0.95	**0.98**	**0.97**	**0.98**	0.95	0.88	0.94	0.96	0.92	0.96

The results indicate that optical water indices, particularly S2-AWEI, consistently achieved higher classification performance than SAR-only features across all evaluated machine learning algorithms. Among all feature combinations, S2-AWEI produced the best overall performance, with the Support Vector Machine (SVM) classifier achieving an overall accuracy (OA) of 0.98, a Kappa coefficient of 0.97, and an F1-score of 0.98, while Random

Forest (RF) achieved an OA of 0.97 with similarly high classification consistency. These findings suggest that AWEI effectively enhances water-related spectral characteristics while suppressing non-water background noise.

In contrast, the SAR-only feature configuration produced lower and less consistent results, particularly for the SVM model, which achieved a recall value of only 0.443 for the water class. This limitation is likely associated with spectral ambiguity in wetland environments, where smooth wet surfaces, saturated soils, and vegetated wetlands may exhibit similar SAR backscatter responses to open water surfaces.

The integration of SAR and optical features generally improved classification stability and produced more balanced precision and recall values across water and non-water classes. In particular, the SAR+FWEI feature combination achieved perfect precision and recall values for several models, indicating strong capability in distinguishing inundated and non-inundated areas. However, integrated features did not consistently outperform the S2-AWEI feature set in terms of overall accuracy and F1-score.

Among the evaluated classifiers, SVM and Logistic Regression demonstrated the most stable performance across multiple feature combinations, while XGBoost achieved competitive results with high precision values. Overall, the experimental results demonstrate that optical water indices play the dominant role in water-body discrimination, whereas SAR backscatter provides complementary structural information that improves classification robustness under heterogeneous wetland conditions.

Table 4. User Accuracy (Precision) for Water and Non-water Classes

Model	**RF**		**SVM**		**LR**		**XGB**	
	1	**0**	**1**	**0**	**1**	**0**	**1**	**0**
SAR	0.942	0.986	0.779	0.999	0.953	0.997	0.930	1.000
S2-FWEI	0.725	0.745	0.894	0.761	0.818	0.793	0.869	0.768
S2-AWEI	0.966	0.966	0.980	0.973	0.979	0.975	0.980	0.971
SAR+FWEI	1.000	0.865	1.000	0.872	1.000	0.851	1.000	0.874
SAR+AWEI	0.906	0.985	0.988	0.984	1.000	0.926	0.945	0.945

Table 5. Producer Accuracy (Recall) for Water and Non-water Classes

Model	**RF**		**SVM**		**LR**		**XGB**	
	1	**0**	**1**	**0**	**1**	**0**	**1**	**0**
SAR	0.885	0.993	0.443	1.000	0.920	0.998	0.862	1.000
S2-FWEI	0.734	0.733	0.688	0.928	0.753	0.840	0.718	0.896
S2-AWEI	0.964	0.967	0.974	0.981	0.977	0.980	0.970	0.981
SAR+FWEI	1.000	0.853	1.000	0.861	1.000	0.834	1.000	0.864
SAR+AWEI	0.973	0.945	0.970	0.994	0.853	1.000	0.892	0.998

Note: Class label 1 represents water, whereas class label 0 represents non-water.

User Accuracy (precision) and Producer Accuracy (recall) were separately evaluated for water (class 1) and non-water (class 0) categories because the classification task focuses on water-body detection. As shown in Tables 4 and 5, the SAR+FWEI feature combination achieved high precision and recall values for the water class across multiple classifiers, indicating strong capability in identifying flooded and water-covered areas. However, its performance for the non-water class remained relatively lower, which contributed to reduced overall accuracy and F1-score compared to the SAR+AWEI feature combination.

The SAR+FWEI feature combination exhibited limitations in distinguishing non-water surfaces, whereas the SAR+AWEI feature combination consistently demonstrated stable and balanced performance across all evaluation metrics. In particular, the SVM classifier achieved the highest overall accuracy (0.98), Kappa coefficient (0.97), and F1-score (0.98), although its performance remained statistically comparable to the Random Forest classifier.

Furthermore, Tables 4 and 5 demonstrate that the SAR+AWEI feature combination maintained relatively balanced precision and recall values for both water and non-water classes. This balanced performance indicates a stronger ability to reduce classification errors under heterogeneous wetland conditions. Overall, the results suggest that integrating SAR backscatter with AWEI provides better generalization capability and more robust classification performance than both single-source features and other hybrid feature combinations.

3.6. Confusion Matrix

Figure 6 presents the confusion matrices for all classifiers using the SAR–AWEI feature combination. Overall, the results demonstrate that all classifiers achieved strong capability in distinguishing water and non-water classes. Nevertheless, differences in false positive (FP) and false negative (FN) distributions indicate variations in model sensitivity and generalization performance.

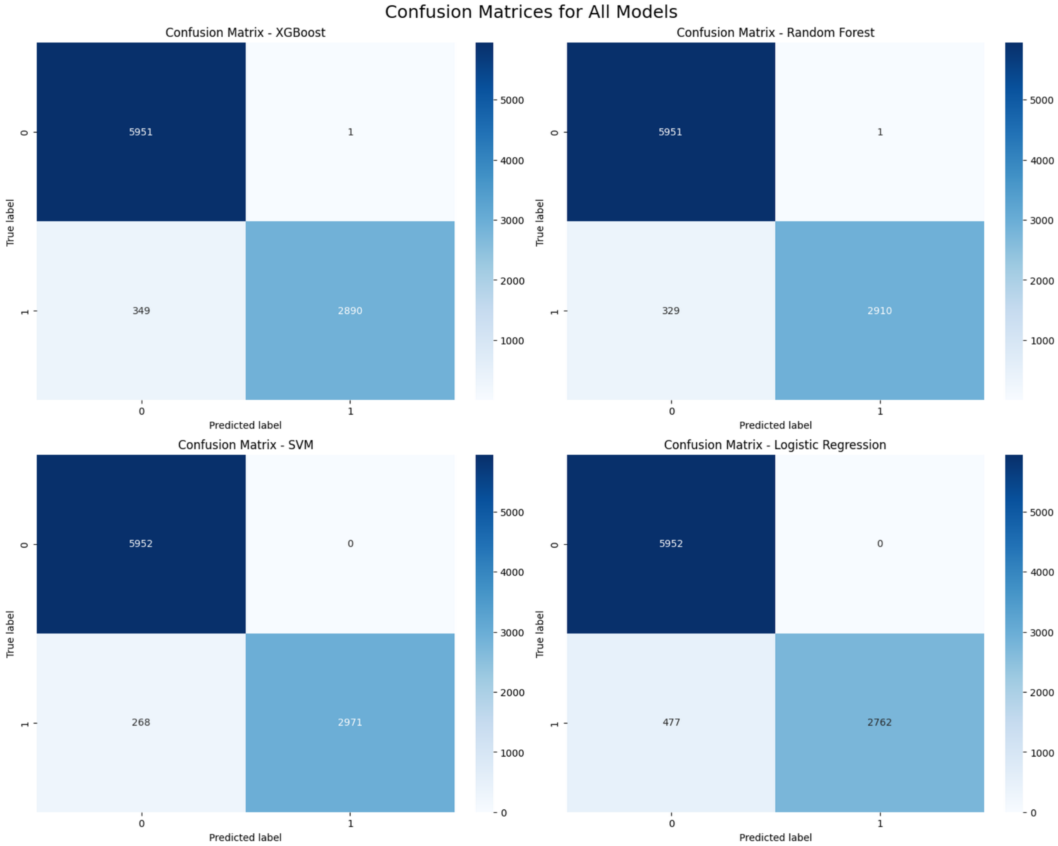


Figure 6. Confusion matrices for RF, SVM, LR, and XGBoost using SAR+AWEI features

Among the evaluated classifiers, the SVM model achieved the most balanced classification performance, with the highest number of true positives (TP = 2971) and true negatives (TN = 5952), while producing no false positives (FP = 0). Although the model still generated a moderate number of false negatives (FN = 268), it outperformed the other classifiers in detecting flooded areas while maintaining high classification precision.

The Random Forest and XGBoost classifiers exhibited similar classification behavior, although both produced slightly higher false negative values (FN = 349). This result suggests that, despite achieving high overall accuracy, their sensitivity to flooded areas was lower than that of the SVM classifier. In contrast, Logistic Regression produced the highest number of false negatives (FN = 477), indicating relatively limited capability in capturing complex nonlinear relationships within the integrated SAR–AWEI feature space.

These findings indicate that all evaluated classifiers achieved strong overall classification performance; however, the SVM classifier provided the best balance between precision and recall by minimizing both false positive and false negative errors. The superior performance of SVM further demonstrates its effectiveness in modeling integrated SAR–optical feature representations for accurate flood detection under heterogeneous wetland conditions.

3.7. Feature Correlation Analysis

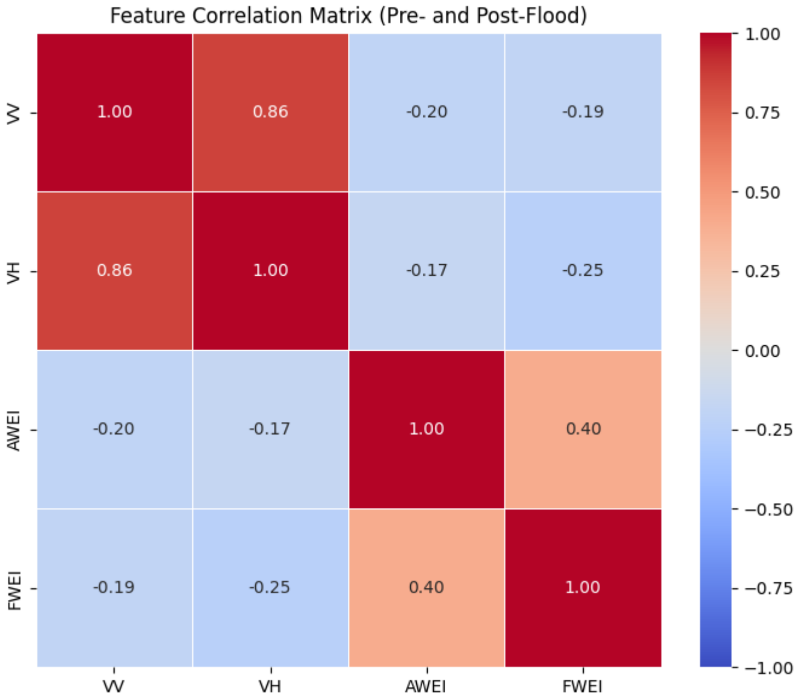


Figure 7. Feature Correlation Matrix of SAR backscatter and optical water indices from pre- and post-flood conditions

Figure 7 presents the Pearson correlation matrix of the SAR backscatter features (VV and VH) and optical water indices (AWEI and FWEI) derived from the combined pre- and post-flood datasets. The results indicate a strong positive correlation between the VV and VH backscatter coefficients ($r = 0.86$), which is expected because both variables originate from the same SAR acquisition system and capture similar surface scattering characteristics. In contrast, the correlation between the optical-derived indices AWEI and FWEI is moderate ($r = 0.40$), indicating that both indices provide related yet non-redundant information for water detection.

The SAR backscatter features also exhibit weak negative correlations with the optical indices, with correlation values ranging from -0.17 to -0.25. This pattern suggests that SAR and optical features capture different physical surface properties and therefore contribute complementary information to the classification process. The relatively low inter-feature correlations further indicate limited multicollinearity within the proposed hybrid SAR–optical method, supporting the robustness and effectiveness of the multi-sensor integration approach for flood inundation mapping.

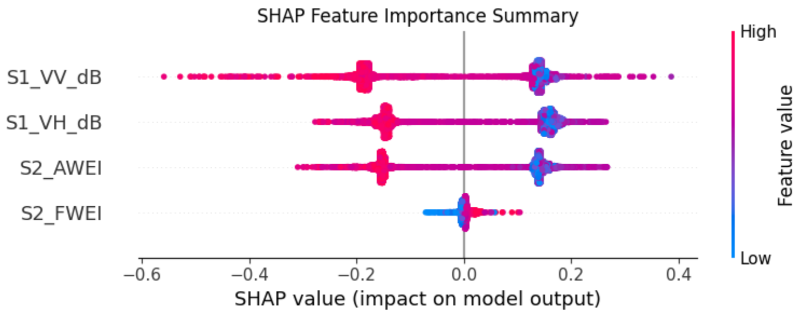


Figure 8. SHAP Feature Importance

3.8. *SHAP / Feature Importance*

Figure 8 presents the SHAP summary plot for the Random Forest model using the four selected features (VV, VH, AWEI, and FWEI). SHAP (Shapley Additive Explanations) analysis was applied to interpret the contribution of individual features to the classification model. SHAP values quantify how each feature influences the model output through positive or negative contributions to the prediction.

Although SAR features exhibited substantial SHAP contributions within the hybrid model, AWEI achieved the strongest standalone classification performance. This finding highlights the complementary nature of the two data sources, where SAR contributes structural and moisture-related information, while AWEI enhances spectral discrimination of water bodies. Among the optical-derived indices, AWEI demonstrated a greater contribution to the model output than FWEI. The SHAP value distributions further suggest that the model captures nonlinear relationships between SAR backscatter responses and optical water characteristics.

The improved performance of the SAR–AWEI feature combination can be explained by the complementary physical characteristics captured by both data sources. SAR backscatter features are sensitive to surface roughness, vegetation structure, and moisture conditions, enabling the detection of inundated surfaces even under partially vegetated wetland environments. Meanwhile, AWEI enhances water detection by emphasizing the spectral absorption characteristics of water bodies while suppressing non-water surfaces.

In heterogeneous wetlands, optical indices alone may incorrectly classify moist soil, shadows, or dense vegetation as flooded areas, resulting in false positives. Conversely, SAR data alone may produce ambiguities in areas with complex vegetation structures and varying surface scattering responses. By integrating SAR backscatter information with the AWEI index, the proposed method improves the separation between flooded and non-flooded surfaces, thereby reducing misclassification and producing more spatially consistent flood inundation maps.

Overall, the SHAP analysis confirms that the classification model does not rely on a single dominant variable but instead benefits from complementary contributions from both SAR and optical features. This finding supports the robustness and interpretability of the proposed hybrid SAR–optical method for flood inundation mapping.

3.9. *Statistical Significance Test*

The Friedman test was performed using the OA values obtained from cross-validation experiments to evaluate the relative performance consistency of the evaluated classifiers across all feature configurations. The average ranks and standard deviations for each classifier are presented in Table 6.

Random Forest achieved the best average rank (1.00), indicating the most consistent classification performance across the evaluated experiments. XGBoost and SVM obtained intermediate average ranks of 2.05 and 3.02, respectively, while Logistic Regression showed the lowest overall ranking (3.93). Although SVM achieved the highest OA and F1-score values for several feature combinations, the Friedman ranking indicates that Random Forest provided more stable performance across cross-validation folds.

To assess whether the observed differences among classifiers were statistically significant, a post-hoc Nemenyi test was performed using the cross-validation results. The analysis involved four classifiers ($k = 4$) evaluated across

Table 6. Friedman Average Ranks and Standard Deviation Across Cross-Validation Folds

Classifier	**Average Rank**	**Std. Dev. Rank**
Random Forest	1.00	0.0000
XGBoost	2.05	0.0612
SVM	3.02	0.0570
Logistic Regression	3.93	0.0671

ten folds ($N = 10$) at a significance level of $\alpha = 0.05$. Based on these parameters, the critical value (q_α) was 2.569, resulting in a critical difference (CD) of 1.4832.

Table 7. Nemenyi Post-hoc Test Results

Model 1	**Model 2**	**Average Rank Difference**	**Significant ($>$ CD)**
XGBoost	Random Forest	1.05	False
XGBoost	SVM	1.00	False
XGBoost	Logistic Regression	1.85	True
Random Forest	SVM	2.05	True
Random Forest	Logistic Regression	2.90	True
SVM	Logistic Regression	0.85	False

The Nemenyi post-hoc test, presented in Table 7, was performed using the OA values obtained from the cross-validation experiments employed in the Friedman analysis. The test evaluates pairwise differences among classifiers based on their average ranking differences across all evaluated feature configurations.

As shown in Table 7, ”True” indicates a statistically significant difference between two classifiers, whereas ”False” indicates no significant difference. Significant differences were observed between Random Forest and SVM, Random Forest and Logistic Regression, as well as XGBoost and Logistic Regression, since their average rank differences exceeded the critical difference threshold.

In contrast, no statistically significant differences were found between XGBoost and Random Forest, XGBoost and SVM, or SVM and Logistic Regression, suggesting relatively comparable classification performance among these model pairs.

Although SVM achieved the highest OA and F1-score values for several feature combinations, the Friedman ranking results indicate that Random Forest provided the most consistent overall performance across the cross-validation experiments. Overall, the statistical analysis demonstrates that both Random Forest and SVM achieved robust and competitive performance for hybrid SAR–optical water classification and flood inundation mapping.

3.10. Water Body Map Using Machine Learning

Figure 9 presents the water body maps generated using the best-performing configuration of SAR and optical water indices (FWEI and AWEI) with the SVM classifier. The model was trained using multiple feature sets, including SAR backscatter, optical water indices, and their integrated combinations. Spatially independent water and non-water samples collected from both pre-flood and post-flood conditions were used for model training and validation. The trained model was subsequently applied separately to the pre-flood and post-flood datasets to generate binary classification maps representing water and non-water classes.

The resulting maps demonstrate that each feature configuration produces distinct spatial patterns of detected water bodies. The SAR-based classification shows a relatively sparse and fragmented distribution of water pixels, particularly under pre-flood conditions, indicating limited sensitivity to smaller or spectrally complex water surfaces. In contrast, the FWEI-based results exhibit a denser distribution of detected water pixels, especially along the river network, although they also introduce higher background noise through the misclassification of non-water surfaces. The AWEI-based classification provides a more balanced representation by clearly delineating the main river channel while reducing background noise compared with the FWEI-based approach.

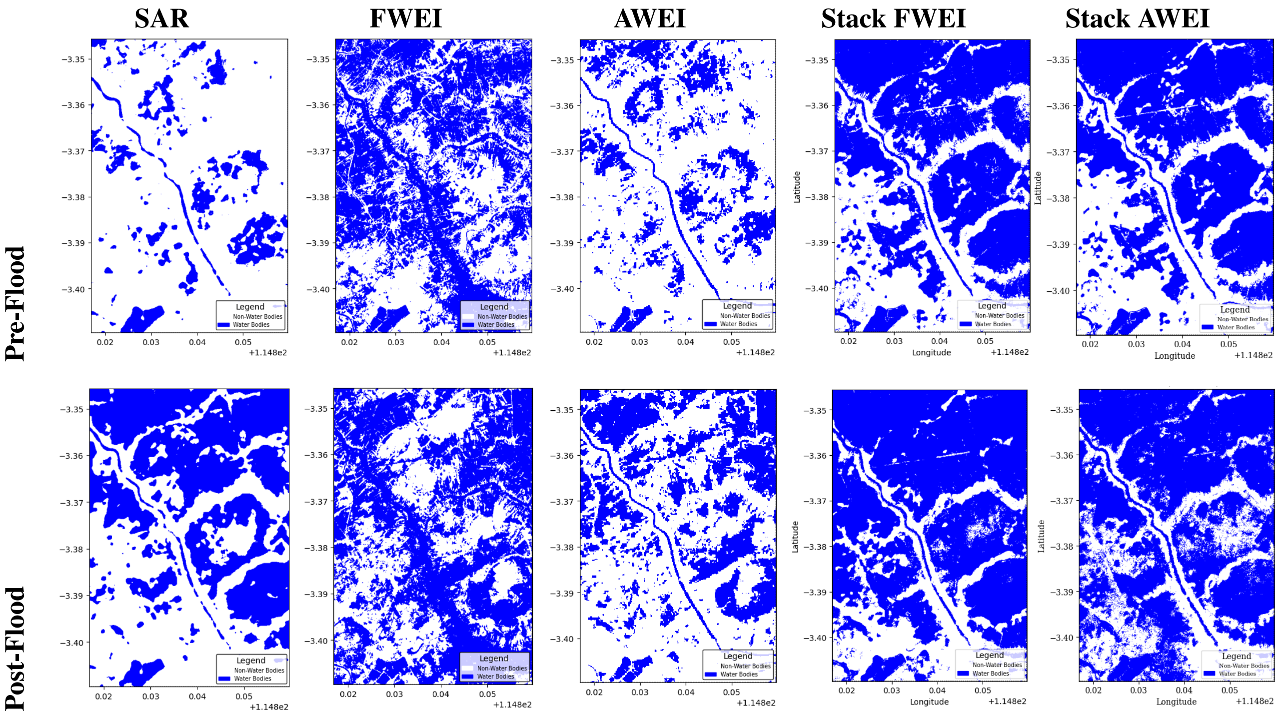


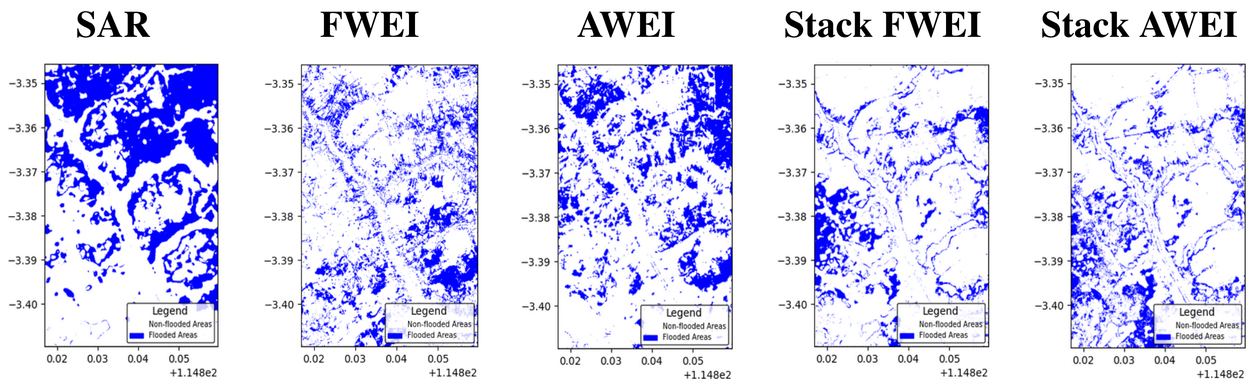
Figure 9. Water Bodies Mapping Generated Using SAR and Optical Water Indices

The integrated feature combinations (SAR+FWEI and SAR+AWEI) produced more spatially coherent and continuous water-body patterns with reduced pixel-level noise relative to single-feature classifications. This improvement is particularly important in wetland environments, where variations in vegetation cover, soil moisture, and backscatter responses often complicate water detection. Nevertheless, several misclassifications remain in densely vegetated and saturated soil areas due to their spectral and backscatter similarity to open water surfaces.

Overall, the results demonstrate that the SVM classifier effectively captures flood-induced surface changes while maintaining spatial consistency across multiple feature combinations. The findings further highlight the importance of SAR–optical feature integration for improving water detection performance in heterogeneous wetland environments. Flood inundation was subsequently identified through temporal comparison between the independently classified pre-flood and post-flood water maps. Pixels that changed from non-water in the pre-flood condition to water in the post-flood condition were classified as inundated areas. The resulting inundation map was further refined using a land-cover mask to remove permanent water bodies, thereby retaining only flood-related inundation areas. The final flood inundation map generated using the best-performing model and feature combination is presented in the following figure.

3.11. Flooded Inundation Map

Figure 10 shows flood inundation maps that were derived by comparing pre- and post-flood water classification using the best-performing hybrid feature (SAR-AWEI). Flooded areas are identified as pixels that transition from non-water bodies in the pre-flood condition to water bodies in the post-flood condition. The resulting maps indicate that all the feature combinations consistently show that flooding is mostly happening along the main river channel and spreading into nearby low-lying floodplain areas. This spatial distribution illustrates the hydrological dynamics of the research area, characterized by river overflow and surface water accumulation in flat wetland ecosystems. The delineated flood zones follow natural drainage routes and topographical depressions, indicating that the model accurately reflects the fundamental flood dynamics.



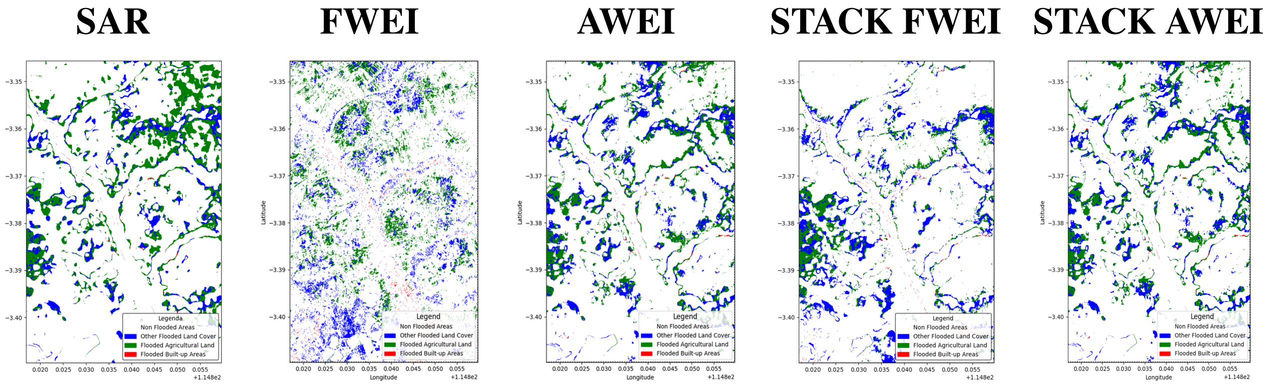
SAR FWEI AWEI Stack FWEI Stack AWEI

Figure 10. Flood Inundation Maps Generated using Best Feature SAR and Optical Water Indices

However, noticeable differences exist among feature configurations in terms of flood extent and classification noise of the detected flood areas. The SAR-based analysis indicates that the flooded pixels are dispersed and fragmented in a manner consistent with the backscatter variability influenced by vegetation, uneven terrain, and soil moisture content. The FWEI-based picture illustrates an extensive area of inundation, particularly surrounding the river corridor. It generates considerable noise in adjacent areas due to its erroneous detection of non-aquatic surfaces.

The result derived using AWEI is superior to that based on FWEI, as it more accurately illustrates flood patterns and exhibits reduced noise. However, it still perceives an excess of water. Conversely, the stacked features, specifically SAR-FWEI and SAR-AWEI, produce inundation patterns distinguished by enhanced spatial coherence and continuity, coupled with diminished pixel-level noise and improved delineation of flood boundaries. These observations imply that the fusion of SAR and optical data augments the dependability of flood detection, especially within wetland environments where data originating from a singular source may be inadequate for accurately depicting intricate surface conditions.

3.12. Flooded Inundation Land Cover Map



SAR FWEI AWEI STACK FWEI STACK AWEI

Figure 11. Results of the Flood-affected Land Cover Analysis Map

Figure 11 presents the flood-affected land-cover maps generated using different SAR and optical feature configurations. The maps illustrate the spatial distribution of flood-related impacts on agricultural and built-up areas derived from temporal comparison between pre-flood and post-flood classifications.

Noticeable differences can be observed among the evaluated feature combinations. The SAR-based result exhibits relatively fragmented and scattered flood-related patches, indicating sensitivity to backscatter variability caused by vegetation cover, surface roughness, and soil moisture conditions. In contrast, the FWEI-based map produces a denser distribution of flooded areas but introduces considerable classification noise through the overestimation of flood-affected regions. The AWEI-based configuration provides a more spatially coherent representation with reduced background noise compared with FWEI, although several overestimated inundated regions remain visible.

The integrated SAR–optical feature combinations, particularly SAR+AWEI, produce more continuous and spatially consistent flood patterns than single-feature approaches. This improvement is especially evident in areas where built-up surfaces, vegetation, and wetland environments are closely intermixed. The results indicate that combining SAR backscatter with optical water indices improves the representation of heterogeneous wetland landscapes and enables more reliable delineation of flood-affected land-cover areas.

Overall, the SVM classifier achieved the highest OA and F1-score values among the evaluated models, while the Nemenyi post-hoc test indicated that its performance was statistically comparable to Random Forest. These findings suggest that both SVM and Random Forest provide robust and competitive performance for hybrid SAR–optical flood inundation and land-cover analysis.

3.13. Flood Inundation Evaluation

To further assess the effectiveness of the proposed method for flood mapping applications, the final flood inundation maps derived from the best-performing machine learning model were evaluated and compared with a conventional Otsu thresholding baseline, as presented in Table 8. Flood inundation maps were generated through post-classification change detection by identifying transitions from non-water in the pre-flood condition to water in the post-flood condition.

Table 8. Flood inundation evaluation results

Method	OA	IoU	Precision	Recall	F1
Hybrid SAR–Optical ML	0.825	0.650	1.00	0.65	0.79
Otsu Thresholding	0.550	0.379	0.55	0.55	0.55

The evaluation results indicate that the machine learning-based flood mapping approach substantially outperformed the threshold-based baseline. Flood inundation assessment was conducted using 200 spatially independent validation samples, consisting of 100 flooded samples (class 1) and 100 non-flooded samples (class 0). The reference labels were derived through visual interpretation of pre-flood and post-flood satellite imagery and subsequently used to evaluate the final flood inundation maps. The proposed hybrid SAR–optical machine learning method achieved an overall accuracy (OA) of 0.825, an Intersection over Union (IoU) value of 0.650, and an F1-score of 0.79. In contrast, the Otsu thresholding approach produced lower performance, with an OA of 0.550, an IoU value of 0.379, and an F1-score of 0.55.

Although both approaches were evaluated using the same validation samples, the threshold-based method exhibited limitations in distinguishing flooded and non-flooded surfaces under heterogeneous wetland conditions. The hybrid SAR–optical machine learning method generated more balanced and spatially consistent flood delineation results by integrating complementary SAR backscatter and optical water-index information.

4. Discussion

The improved performance of the proposed method can be attributed to the complementary characteristics of SAR backscatter and optical water indices. SAR data provides sensitivity to surface roughness, moisture conditions, and vegetation structure, while optical indices enhance water-related spectral absorption characteristics. Their integration within a supervised machine learning method enables more effective discrimination of complex inundation patterns compared to conventional threshold-based approaches.

The superior performance of the hybrid SAR–optical method can be attributed to the complementary characteristics of radar and optical observations. While SAR data capture surface roughness and moisture-related variations regardless of cloud cover, optical water indices provide strong spectral discrimination of surface water. The integration of these complementary features is particularly suitable for flood mapping in heterogeneous wetland environments where single-sensor approaches may be affected by spectral ambiguity, vegetation cover, or moisture-induced backscatter variability.

Although deep learning approaches have become increasingly popular in remote sensing applications by 2026, traditional machine learning methods remain highly relevant for operational flood response applications because of their lower computational requirements, faster training processes, and improved interpretability under limited training data conditions. This study demonstrates that the proposed hybrid SAR–optical machine learning method can achieve reliable flood inundation mapping performance without relying on computationally intensive deep learning architectures.

5. Conclusion

This study developed and evaluated a hybrid SAR–optical method for flood inundation mapping in heterogeneous wetland environments. First, the proposed pixel-wise feature stacking approach integrates Sentinel-1 SAR backscatter and Sentinel-2 optical water indices, enabling improved discrimination between water and non-water classes under complex surface conditions. The experimental results demonstrate that multi-sensor integration, particularly the SAR–AWEI feature combination, produced higher classification accuracy and more spatially coherent flood delineation results compared to single-source features. The findings are consistent with previous studies demonstrating that SAR–optical integration improves flood delineation by combining radar-derived structural information with the strong spectral sensitivity of optical data to surface water. This synergy enables more accurate discrimination between flooded and non-flooded areas, particularly in heterogeneous wetland environments [28, 29].

Second, the evaluation of multiple machine learning algorithms revealed that the Support Vector Machine (SVM) classifier achieved the highest overall accuracy and F1-score values, although its performance was statistically comparable to Random Forest. This finding highlights the importance of selecting appropriate classifiers for handling high-dimensional and heterogeneous remote sensing data in flood mapping applications.

Third, feature contribution analysis using SHAP combined with spatially independent validation provided important insights into model robustness and generalization capability. The results indicate that AWEI contributed substantially to water discrimination, while SAR backscatter features provided complementary structural information that improved classification stability under heterogeneous wetland conditions. Furthermore, the use of spatially independent validation reduced bias caused by spatial autocorrelation and enabled a more realistic assessment of model performance.

It is important to note that the flood inundation maps generated in this study were derived from pre- and post-classification change detection analysis. Consequently, the accuracy of flood inundation mapping is inherently dependent on the reliability of water classification results and may be affected by cumulative classification errors between temporal stages.

Despite the promising results, several limitations should be acknowledged. First, the study was conducted within a single study area (Banjar Regency), which may limit the generalizability of the method to regions with different geomorphological and hydrological characteristics. Second, misclassification may still occur in areas with dense vegetation and water-saturated soils because of similar spectral and SAR backscatter responses. Third, the analysis relied on bi-temporal composite imagery representing pre-flood and post-flood periods. Although this approach reduces cloud contamination and improves data availability, it may not fully capture the exact temporal dynamics of individual flood events and may introduce uncertainty in distinguishing temporary flood inundation from seasonal water variability.

Future research may explore the integration of deep learning approaches and multi-temporal datasets to improve scalability, temporal sensitivity, and robustness across diverse geographic regions. Nevertheless, the present study demonstrates that conventional machine learning approaches remain effective, interpretable, and computationally efficient for operational flood inundation mapping in heterogeneous wetland environments.

Acknowledgement

We gratefully acknowledge the Geosoftware Community (GeosoftwareID) for their valuable resources and webinars on Google Earth Engine (GEE), which supported this research.

Author Contributions

Conceptualization, methodology, software development, validation, formal analysis, investigation, and writing—original draft preparation, D.I.P.; writing—review and editing, C.S.O.; supervision, E.N.S., P.R.W., and A.R.S. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

REFERENCES

1. R. A. Balgah, K. A. Ngwa, G. R. Buchenrieder, and J. N. Kimengsi, "Impacts of Floods on Agriculture-Dependent Livelihoods in Sub-Saharan Africa: An Assessment from Multiple Geo-Ecological Zones," *Land*, vol. 12, no. 2, p. 334, Feb. 2023. [Online]. Available: https://www.mdpi.com/2073-445X/12/2/334
2. A. Voitik, V. Kravchenko, O. Pushka, T. Kutkovetska, T. Shchur, and S. Kocira, "Comparison of NDVI, NDRE, MSAVI and NDSI Indices for Early Diagnosis of Crop Problems," *Agricultural Engineering*, vol. 27, no. 1, pp. 47–57, Jan. 2023. [Online]. Available: https://reference-global.com/article/10.2478/agriceng-2023-0004
3. M. Zhang, Z. Chen, J. Wang, B. Kar, M. Pierce, K. Tiampo, R. Eguchi, and M. Glasscoe, "Optical Remote Sensing for Global Flood Disaster Mapping: A Critical Review Towards Operational Readiness," *Remote Sensing*, vol. 17, no. 11, p. 1886, Jan. 2025. [Online]. Available: https://www.mdpi.com/2072-4292/17/11/1886
4. E. Eva, "BPBD Banjar: 17.257 Rumah di 7 Kecamatan Terendam Banjir," Mar. 2023. [Online]. Available: https://news.detik.com/berita/d-6601374/bpbd-banjar-17-257-rumah-di-7-kecamatan-terendam-banjir
5. Hasanuddin, "Banjir 2023, BPBD Kalsel; Puluhan Ribu Jiwa Terdampak," Apr. 2023. [Online]. Available: https://bakabar.com/post/banjir-2023-bpbd-kalsel-puluhan-ribu-jiwa-terdampak-1g4m40e3
6. A. Muhari, "Badan Nasional Penanggulangan Bencana," Feb. 2023. [Online]. Available: https://www.bnpb.go.id/berita/banjir-genangi-ribuan-rumah-di-kabupaten-banjar-kalimantan-selatan
7. F. Lestari, M. K. Sudaryo, R. Djalante, A. Adiwibowo, A. Kadir, Zakianis, and S. A. Satyawardhani, "Estimating the Flood, Landslide, and Heavy Rainfall Susceptibility of Vaccine Transportation after 2021 Flooding in South Kalimantan Province, Indonesia," *Sustainability*, vol. 16, no. 4, p. 1554, Jan. 2024. [Online]. Available: https://www.mdpi.com/2071-1050/16/4/1554
8. J. A. Rahma, S. Ratna, M. Muflih, M. Amin, D. I. Puspitasari, and H. Budiman, "Time Series Prediction Of Martapura River Water Level Utilising Ensemble Stacked LSTM, BPNN, Random Forest And XGBoost," in *2025 7th International Conference on Cybernetics and Intelligent System (ICORIS)*, Sep. 2025, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11295996

9. D. I. Puspitasari, E. Noersasongko, M. A. Soeleman, Purwanto, and P. N. Andono, “Machine Learning Algorithms for Land Cover Classification Using Hybrid Remote Sensing and Spectral Indices,” in *2024 International Seminar on Application for Technology of Information and Communication (iSemantic).* IEEE, Sep. 2024, pp. 377–381. [Online]. Available: https://ieeexplore.ieee.org/document/10762048/
10. C. Zhao, H. Wei, G. L. Feyisa, T. De Castro Tayer, G. Ma, H. Wu, and Y. Pan, “Evaluating spectral indices for water extraction: Limitations and contextual usage recommendations,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 139, p. 104510, May 2025. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1569843225001578
11. H. Farhadi, H. Ebadi, A. Kiani, and A. Asgary, “A novel flood/water extraction index (FWEI) for identifying water and flooded areas using sentinel-2 visible and near-infrared spectral bands,” *Stochastic Environmental Research and Risk Assessment*, vol. 38, no. 5, pp. 1873–1895, May 2024. [Online]. Available: https://link.springer.com/10.1007/s00477-024-02660-z
12. L. Che, S. Li, and X. Liu, “Improved surface water mapping using satellite remote sensing imagery based on optimization of the Otsu threshold and effective selection of remote-sensing water index,” *Journal of Hydrology*, vol. 654, p. 132771, Jun. 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S002216942500109X
13. X. Peng, S. Chen, Z. Miao, Y. Xu, M. Ye, and P. Lu, “Automatic Flood Monitoring Method with SAR and Optical Data Using Google Earth Engine,” *Water*, vol. 17, no. 2, p. 177, Jan. 2025. [Online]. Available: https://www.mdpi.com/2073-4441/17/2/177
14. R. Colacicco, A. Refice, R. Nutricato, F. Bovenga, G. Caporusso, A. D’Addabbo, M. La Salandra, F. P. Lovergine, D. O. Nitti, and D. Capolongo, “High-Resolution Flood Monitoring Based on Advanced Statistical Modeling of Sentinel-1 Multi-Temporal Stacks,” *Remote Sensing*, vol. 16, no. 2, p. 294, Jan. 2024. [Online]. Available: https://www.mdpi.com/2072-4292/16/2/294
15. E. Portalés-Julià, G. Mateo-García, and L. Gómez-Chova, “Understanding Flood Detection Models Across Sentinel-1 and Sentinel-2 Modalities and Benchmark Datasets,” 2025. [Online]. Available: https://www.ssrn.com/abstract=5118486
16. J. Deppner and M. Cajias, “Accounting for Spatial Autocorrelation in Algorithm-Driven Hedonic Models: A Spatial Cross-Validation Approach,” *The Journal of Real Estate Finance and Economics*, vol. 68, no. 2, pp. 235–273, Feb. 2024. [Online]. Available: https://link.springer.com/10.1007/s11146-022-09915-y
17. D. Koldasbayeva, P. Tregubova, M. Gasanov, A. Zaytsev, A. Petrovskaia, and E. Burnaev, “Challenges in data-driven geospatial modeling for environmental research and practice,” *Nature Communications*, vol. 15, no. 1, p. 10700, Dec. 2024. [Online]. Available: https://www.nature.com/articles/s41467-024-55240-8
18. F. Foroughnia, S. M. Alfieri, M. Menenti, and R. Lindenbergh, “Evaluation of SAR and Optical Data for Flood Delineation Using Supervised and Unsupervised Classification,” *Remote Sensing*, vol. 14, no. 15, Aug. 2022. [Online]. Available: https://www.mdpi.com/2072-4292/14/15/3718
19. H. Ren, B. Pang, P. Bai, G. Zhao, S. Liu, Y. Liu, and M. Li, “Flood Susceptibility Assessment with Random Sampling Strategy in Ensemble Learning (RF and XGBoost),” *Remote Sensing*, vol. 16, no. 2, p. 320, Jan. 2024. [Online]. Available: https://www.mdpi.com/2072-4292/16/2/320
20. J. Tong, F. Gao, H. Liu, J. Huang, G. Liu, H. Zhang, and Q. Duan, “A Study on Identification of Urban Waterlogging Risk Factors Based on Satellite Image Semantic Segmentation and XGBoost,” *Sustainability*, vol. 15, no. 8, p. 6434, Jan. 2023. [Online]. Available: https://www.mdpi.com/2071-1050/15/8/6434
21. M. Zhao and N. Ye, “High-Dimensional Ensemble Learning Classification: An Ensemble Learning Classification Algorithm Based on High-Dimensional Feature Space Reconstruction,” *Applied Sciences*, vol. 14, no. 5, Feb. 2024. [Online]. Available: https://www.mdpi.com/2076-3417/14/5/1956
22. Y. Wu, Z. Zhang, X. Qi, W. Hu, and S. Si, “Prediction of flood sensitivity based on Logistic Regression, eXtreme Gradient Boosting, and Random Forest modeling methods \textbar Water Science & Technology \textbar IWA Publishing,” *Water Science & Technology*, vol. 89, no. No, 10, May 2024. [Online]. Available: https://iwaponline.com/wst/article/89/10/2605/102133/Prediction-of-flood-sensitivity-based-on-Logistic
23. C. I. Cîmpianu, A. Mihu-Pintilie, C. C. Stoleriu, A. Urzică, and E. Huţanu, “Managing Flood Hazard in a Complex Cross-Border Region Using Sentinel-1 SAR and Sentinel-2 Optical Data: A Case Study from Prut River Basin (NE Romania),” *Remote Sensing*, vol. 13, no. 23, p. 4934, Jan. 2021. [Online]. Available: https://www.mdpi.com/2072-4292/13/23/4934
24. N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-9, no. 1, Jan. 1979.
25. M. Moharrami, M. Javanbakht, and S. Attarchi, “Automatic flood detection using sentinel-1 images on the google earth engine,” *Environmental Monitoring and Assessment*, vol. 193, no. 5, p. 248, Apr. 2021. [Online]. Available: https://doi.org/10.1007/s10661-021-09037-7
26. M. G. Gümüş, “Performance Analysis of Water Extraction Indices with Geospatial and Statistical Techniques Using Google Earth Engine Platform: A Case Study of Ramsar Wetlands in Türkiye,” *Journal of the Indian Society of Remote Sensing*, vol. 53, no. 8, pp. 2697–2721, Aug. 2025. [Online]. Available: https://link.springer.com/10.1007/s12524-025-02180-5
27. M. Tesfaye and L. Breuer, “Performance of water indices for large-scale water resources monitoring using Sentinel-2 data in Ethiopia,” *Environmental Monitoring and Assessment*, vol. 196, no. 5, p. 467, May 2024. [Online]. Available: https://link.springer.com/10.1007/s10661-024-12630-1
28. M. Fawakherji and L. Hashemi-Beni, “Flood detection and mapping through multi-resolution sensor fusion: integrating UAV optical imagery and satellite SAR data,” *Geomatics, Natural Hazards and Risk*, vol. 16, no. 1, p. 2493225, Dec. 2025. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/19475705.2025.2493225
29. C. Campo, P. Tamagnone, S. Choy, T. D. Tran, G. J. Schumann, and Y. Kuleshov, “Monitoring Flood Inundation Dynamics From Space,” *Reviews of Geophysics*, vol. 64, no. 1, p. e2025RG000885, Mar. 2026. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2025RG000885